NBER WORKING PAPER SERIES

ECONOMIC ANALYSIS OF ACCIDENT LAW

Steven Shavell

Working Paper 9483
http://www.nber.org/papers/w9483

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2003

ABSTRACT

Accident law is the body of legal rules governing the ability of victims of harm to sue and to collect

payments from those who injured them. This paper contains the chapters on accident law from a

general, forthcoming book, Foundations of Economic Analysis of Law (Harvard University Press,

2003). The analysis is first concerned (chapters 2-4) with the influence of liability rules on

incentives to reduce accident risks. Then consideration of accident law is broadened (chapter 5) to

reflect the effect of liability rules on compensation of victims and the allocation of risk. In this

regard a central issue is the roles of victims' insurance and of liability insurance, and how they alter

the incentives inherent in liability rules. Finally, the administrative costs of the liability system,

namely, the private and public legal costs of litigation, are examined (chapter 6). These costs are

significant and thus bear importantly on whether use of accident law is socially desirable. It is

emphasized that social intervention -- either to curtail use of the legal system or to encourage it --

may well be needed because the private incentives to use the system are generally different from the

socially desirable incentives to do so.

Steven Shavell
Harvard Law School
1575 Massachusetts Avenue
Hauser Hall 508
Cambridge, MA 02138
and NBER
shavell@law.stanford.edu

Table of Contents

Economic Analysis of Accident Law

(part of *Foundations of Economic Analysis of Law*)

Summary Table of Contents of

Foundations of Economic Analysis of Law

(forthcoming 2003, *Harvard University Press*)

Economic Analysis of Accident Law

(part of *Foundations of Economic Analysis of Law*)

Steven Shavell

Chapter 1

INTRODUCTION

1. The Economic Approach to the Analysis of Law

Under the economic approach to the analysis of law, two basic types of questions about legal rules are addressed. The first type is *descriptive*, concerning the effects of legal rules. For example, what is the influence of our system of liability for automobile accidents on the number of these accidents, on the compensation of accident victims, and on litigation expenses? The other type of question is *normative*, pertaining to the social desirability of legal rules. Thus, it might be asked whether our system of liability for automobile accidents is socially good, given its various consequences.

In answering the two types of questions under the economic approach, theoretical attention is usually focused on stylized models of individual behavior and of the legal system. The advantage of studying models is that they allow descriptive and normative questions to be answered in an unambiguous way, and that they may clarify understanding of the actual influence of legal rules on behavior and help in the making of legal policy decisions.

Descriptive analysis. When considering the descriptive questions in the models, the view taken will generally be that actors are forward looking and "rational." That is, they will behave so as to maximize their expected utility. Given the characterization of individuals' behavior as rational, the influence of legal rules on behavior can be ascertained. This can be done with definitude in the world of the models, because all relevant assumptions about individuals' desires, their knowledge, their capabilities, and the environment will have been made explicit. For example, whether a person will drive carefully will be determinable, for it will have been stated how difficult it is for the person to exercise precaution, whether the person will himself be at risk of injury from an automobile accident, what the rule of liability is, what circumstances will give rise to suit, whether the person owns accident insurance and liability insurance, and so forth.

Normative analysis.[1] The evaluation of social policies, and thus of legal rules, will be

---

[1] The discussion to follow here is amplified in chapters 26-28 of the book.

undertaken with reference to a stated measure of *social welfare.* One legal rule will be said to be superior to a second if the first rule results in a higher level of the stated measure of social welfare.

It should therefore be noticed that normative analysis is *conditional* in nature, in that the legal rule that is best may depend on the social welfare criterion under consideration. If the social goal were simply to minimize the number of automobile accidents, the best rule might well involve severe punishment for causing an accident, whereas if the social goal were also to involve the benefits people obtain from driving, the best rule would be unlikely to involve very rigorous punishment for causing an accident.

This raises the question of which measures of social welfare are considered in economic analysis of legal rules. Before discussing this question, let me comment briefly on the nature of the measures of social welfare.

According to the framework of welfare economics, social welfare is assumed to be a function of individuals' well-being, that is, of their utilities. An individual's utility, in turn, can depend on anything about which the individual cares: not only material wants, but also, for example, aesthetic tastes, altruistic feelings, or a desire for notions of fairness to be satisfied. Hence, social welfare can depend on any of these elements, and will depend on them to the extent that individuals' utilities do. It is thus a mistake to believe that, under the economic view, social welfare reflects only narrowly "economic" factors, namely, the amounts of goods and services produced and enjoyed.

Moreover, the measure of social welfare can embody the desirability of equality of utility among individuals. That is, the function that represents the measure of social welfare may be such that it is higher if individuals have similar utilities than if their utilities are dispersed (but have the same sum). Social welfare functions with the property that equality of utilities tends to raise social welfare are studied by economists in many domains, and notably in the area of income taxation.

Therefore, the conception of social welfare employed in welfare economics is quite general and plausibly can accommodate the views of the social good of most readers.

Nevertheless, it is standard for economic analysts to restrict attention to fairly simple measures of social welfare, and I will do that here. Two types of assumptions that I will tend to make should be noted at the outset. One is that the measure of social welfare will usually not accord importance to the distribution of utilities; thus, the effect of legal rules on the distribution of well-being will not be relevant to their evaluation under the measure of social welfare. This assumption is not made because of an opinion that the distribution of utilities is in fact unimportant. Rather, taking the effect of legal rules on distributional factors into account would complicate our analysis and yet would not in the end alter our conclusions. Why would taking into account distributional factors not alter our conclusions? The answer is that society has an income tax and transfer system that it can utilize to redistribute income. Thus, if I were to incorporate the income tax and transfer system into our analysis, a change in that system could offset any undesirable distributional consequence of a legal rule. If, for example, some legal rule turned out on balance to help the rich and hurt the poor, the former could be taxed more heavily and the latter less, so that the use of the rule would not necessarily have any distributional impact

after the optimal adjustment in the tax and transfer system were made. Thus, if one assumes that the income tax and transfer system will be used to effect desirable changes in the distribution of income, the distributional impact of the choice of legal rules should not matter. Of course, one might not make the assumption that the income tax and transfer system would always be used to redistribute wealth beneficially, in which case the choice of legal rules might be decided in part on the basis of their redistributive effects. The other type of simplification that I want to note now concerns notions of fairness and morality. Consider, for example, the classical conception of corrective justice, demanding that a wrongdoer compensate his victim for harm sustained. It is clear, I think, that the idea of corrective justice has substantial importance to individuals, especially in the context of accidents, and thus might be thought to enter into measures of social welfare. I will usually exclude such notions of fairness, however, from the analysis proper for analytical convenience. I will, though, sometimes mention, after the analysis of a legal rule, how a relevant notion of fairness would affect my conclusions. (Also, as I note below, I will take up the general issue of the integration of morality and notions of fairness into normative analysis later in the book.)

What distinguishes economic analysis of law from other analysis of law? One might ask whether there is any qualitative difference between economic analysis of law, as defined here, and other approaches to its assessment. Is it not of interest to every legal analyst to determine how legal rules affect behavior and then to evaluate the rules with reference to some criterion of the social good? The answer would seem to be "yes," and thus in this general sense, one cannot distinguish economic analysis from other analysis of law.

What does seem to mark economic analysis are three characteristics. First, economic analysis emphasizes the use of stylized models and of statistical, empirical tests of theory, whereas other approaches usually do neither. Second, in describing behavior, economic analysis gives much greater weight than other approaches to the view that actors are rational, acting with a view toward the possible consequences of their choices. And third, in normative evaluation, economic analysis makes explicit the measure of social welfare considered, whereas other approaches often leave the criterion of the social good unclear or substantially implicit.

History of the economic approach. The economic view may be said to have originated mainly with writings on crime by Becarria (1767) and, especially, Bentham (1789). Bentham developed in significant detail the idea that legal sanctions may discourage bad conduct and that sanctions should be employed when they will effectively deter but not when they will fail to do so (as with the insane). Curiously, however, after Bentham, the economic approach to law lay largely dormant until the 1960s and 1970s. In that period, Coase (1960) wrote a provocative article on the incentives to reduce harms to neighbors engendered by property rights assignments; Becker (1968) authored an influential article on crime, casting into modern terms and extending Bentham's earlier contributions; Calabresi (1970) published an extended treatment of liability rules and the accident problem; Posner (1972a) wrote a comprehensive textbook and a number of articles, and he established the *Journal of Legal Studies,* in which scholarship in economic analysis of law could be regularly published. Since that time, economic analysis of law has grown fairly rapidly.

2. Outline and Goal

This book is divided into a number of parts, the first of which are concerned with the basic areas of private law -- liability for accidents, contract law, and property law -- and with civil litigation. These areas of law are said to be private because they are enforced by the bringing of suits by private parties.

In the next part of the book, I deal with public law enforcement, in which the state uses enforcement agents, such as police, safety inspectors, and tax auditors, to detect violators and to sanction them. Crime and criminal law are discussed in this part. In a subsequent, brief part of the book, I examine general questions about the overall structure of the legal system, such as why some behavior is controlled through private law and other behavior through public law enforcement.

Then, in the final part, I consider the relationship between welfare economics and morality, issues relating to income distributional equity and the law, and commonly-encountered questions about the economic approach to analysis of law.

The goal of the book is to set out the major elements of economic analysis of the central areas of law in a way that, as I mentioned in the preface, should be accessible to a wide audience. Thus, although most sections of the book are organized around models of behavior and of the legal system, the analysis is not technical in nature (all mathematics is contained in footnotes). Accordingly, legal readers without any formal background in economics should find the book easy to understand. At the same time, economists should find the subject matter to be of natural intellectual appeal, and they should not have any difficulty owing to their lack of legal background, for the law under discussion is for the most part basic, common knowledge or close to that.

Chapter 2

LIABILITY AND DETERRENCE: BASIC THEORY

Here and throughout Part One, we will be considering a model of accidents involving two types of parties, injurers and victims. We might think, for example, of injurers as drivers of automobiles and of victims as bicyclists, or of injurers as parties conducting blasting operations and of victims as passersby.[2]

Injurers may face legal liability for accidents that they cause, and the effect of this possibility on their behavior, victims' behavior, and specified measures of social welfare will now be considered in several increasingly general versions of the model of accidents. The two major rules of accident liability, strict liability and negligence (and certain variations of them), will be the focus of our analysis.[3]

We will assume that accidents and consequent liability arise probabilistically. In order to analyze the effects of liability rules in an uncertain setting in the simplest way, we will often suppose that parties are *risk-neutral*. A risk-neutral party makes decisions on the basis of probability-discounted, or *expected*, values. For example, a risk-neutral person who faces a liability of $100,000 with probability 10 percent will consider this uncertain payment to be equivalent to a certain payment of its expected value of $10,000.[4] An interpretation of the $10,000 expected liability amount is that it is the payment that the person would make on average were he repeatedly to face a 10 percent risk of having to make a $100,000 payment.

---

[2]Accidents involving parties of only one type – such as accidents involving just drivers of automobiles, or just hunters -- are not in strict logic described by this model. But it will be evident to the reader that many of the conclusions that will be drawn would carry over to a model of these single-activity accidents.

[3]This chapter is based on a more complete treatment of its subject matter in chapter 2 of Shavell 1987a, which also presents proofs of claims that are made. In footnotes to this chapter, however, proofs of a number of the more important conclusions are given or are sketched.

[4]If this assumption were not made, and account were taken of risk aversion, then a liability of $100,000 with 10 percent probability would deter more than a certain liability of $10,000. We will consider risk aversion in chapter 5 below.

1. Unilateral Accidents and Levels of Care

In the first version of the accident model, it will be supposed that accidents are *unilateral* in nature: Only injurers' exercise of care or precautions affects accident risks; victims' behavior does not. When an airplane crashes into a building, for example, or when a rupture in a water main causes a flood in a basement, the victims probably could not have done much to prevent harm. In these cases, the accidents may be seen as almost literally unilateral. Other types of accidents might be seen as approximately unilateral if the victim's role was slight; consider for example automobile-bicycle accidents in which bicyclists' actions are of minor importance in reducing risks.

The social goal will be taken to be minimization of the sum of the costs of care and of expected accident losses. This sum will be called *total social costs*.

1.1 Social welfare optimum. Before determining how injurers will act under different liability rules, let us identify the level of care that minimizes total social costs. This socially optimal level of care will clearly reflect both the costs of exercising care and the reduction in accident risks that care would accomplish. Consider the following example.

Example 1. Suppose that accidents that cause losses of 100 occur with a probability as described in Table 1.

Table 1: Care of Injurers and Accident Risk				
Care level	Cost of care	Probability of accident	Expected accident losses	Total social costs
None	0	15%	15	15
Moderate	3	10%	10	13
High	6	8%	8	14

To understand why exercising moderate care minimizes total social costs, observe on one hand that raising the level of care from none to moderate reduces expected accident losses by 5, but involves costs of only 3; it thus lowers total social costs. On the other hand, raising care beyond the moderate level would reduce expected accident losses by only 2, yet involve additional costs of 3; hence it would not be worthwhile.//

Note that the example illustrates the obvious point that the optimal level of care may well not result in the lowest possible level of expected accident losses (for that would require the highest level of care).[5] Let us now examine how much care injurers will be led to exercise in the absence of liability and under various liability rules.

---

[5]The formal version of the model illustrated in the example is as follows. Let x be the level of care, $p(x)$ the probability of an accident (where p is decreasing in x), and h the harm that an accident would cause. The socially optimal x miminizes $x + p(x)h$ and is denoted by x^*. Unless indicated otherwise, I will assume that x^* is unique in these notes.

1.2 No liability. If there is no liability for accidents, injurers will not exercise any care, for doing so would entail costs but not yield a benefit to them. Total social costs will therefore generally exceed the optimal level; in Example 1, for instance, total social costs will be 15 rather than 13.

1.3 Strict liability. Under the rule of *strict liability*, injurers must, by definition, pay for all accident losses that they cause.[6] Hence, injurers' total costs will equal total social costs; and because injurers will seek to minimize their own total costs, injurers' goal will be identical to the social goal of minimizing total social costs. Consequently, strict liability induces injurers to choose the socially optimal level of care. In Example 1, strict liability leads injurers to exercise the optimal, moderate level of care.

1.4 Negligence rule. Under the *negligence rule*, an injurer is held liable for the accident losses he causes only if he was negligent, that is, only if his level of care was less than a level called *due care* that the courts specify. If the injurer exercised a level of care that equaled or exceeded due care, he will not be held liable. The negligence rule is sometimes said to be *fault-based* because liability is found only if the injurer was at fault in the sense of having been found negligent.

If the courts set the level of due care equal to the socially optimal level of care, then injurers will be led to exercise due care, and thus the outcome will be socially optimal. To see why, first reconsider Example 1. If courts define due care to be the socially optimal, moderate level, the expected liability for an injurer would equal total social costs when no care is taken and would be zero when moderate or high care is taken. When at least moderate care is taken, then, the injurer's total costs equal just the cost of care -- see Table 2. Hence, injurers will indeed be best off exercising moderate care.

Table 2: Negligence Rule				
Care level	Cost of care	Liability	Expected liability	Injurer's total costs
None	0	Yes	15	15
Moderate (due care)	3	No	0	3
High	6	No	0	6

More generally, there are two reasons why injurers will necessarily be led to take due care if it is chosen by courts to equal the optimal level. First, injurers plainly would not take more than due care, because they will escape liability by taking merely due care. Taking greater

---

[6] It is assumed for the most part in this and the next chapter that an injurer is able to pay for losses caused. The important possibility that injurers are unable to pay for losses caused is considered in section 3 of chapter 4.

care would therefore be to no advantage yet would involve additional costs.[7] Second, injurers would not wish to take less than due care if due care is set at the socially optimal level. If injurers took less than due care, they would be exposed to the risk of liability, so their expected costs would equal total social costs. Thus, injurers would want to choose their level of care so as to minimize total social costs. But this in turn means that they would wish to raise their level of care to the socially optimal point -- which by hypothesis equals due care and therefore allows them to avoid liability entirely.[8]

      1.5 Liability rules compared. Both forms of liability result in the same, socially optimal behavior, but they differ in terms of what courts need to know to apply them.[9] Under strict liability a court need only determine the magnitude of the loss that occurred, whereas under the negligence rule a court must in addition determine the level of care actually taken (a driver's speed) and calculate the socially optimal level of due care (the appropriately safe speed). To do the latter, in turn, a court needs to know the costs and the effectiveness of taking different levels of care in reducing accident risks.[10]

      1.6 Several dimensions of care. Suppose, as would be usual, that there is more than one dimension of an injurer's behavior that affects accident risks (not only a driver's speed, but also the frequency with which he looks at the rear-view mirror). In this situation, under strict liability an injurer would be led to choose optimal levels of *all* dimensions of care, because his goal would be to minimize his expected total costs. But under the negligence rule, an injurer would have a motive to choose optimal levels *only* of those dimensions of care that are incorporated in the due care standard. And in fact some dimensions of care will usually be omitted from the due care standard because of difficulties that courts would face in ascertaining them (how would a court obtain information about the number of times per minute a driver usually looks in his rear-view mirror?) or in determining proper behavior in respect to them.

---

    [7]It is assumed here (and elsewhere in this chapter) that a court can determine a party's level of care with complete accuracy. Otherwise, it might well be worth a party's while to take more than due care to reduce the likelihood of a court mistakenly finding him negligent. This and related issues are analyzed in section 1 of chapter 4.

    [8]In terms of the model mentioned in note 4 above, the claim of this section is that if the due care level equals x^*, then injurers will be induced to choose x^*. To demonstrate this, observe that, as stated, an injurer will not choose $x > x^*$, for if he chooses x^*, he spends less and still bears no liability. Thus, $x \leq x^*$ must be true. If $x < x^*$, the injurer will be found negligent if he causes an accident, so that he will bear liability. Thus, given that $x < x^*$, the injurer will choose x to minimize $x + p(x)h$. But $x + p(x)h > x^* + p(x^*)h$ by definition of x^*, and because $x^* + p(x^*)h \geq x^*$, it follows that $x + p(x)h > x^*$; thus, the injurer will prefer to choose x^* than any x less than x^*.

    [9]The rules also differ in how they allocate risk, in the administrative costs that they generate, and in their distributional effects. As indicated in the introduction to Part One, these issues will be discussed in later chapters.

    [10]These disadvantages of the negligence rule (as well as the disadvantage to be noted in the next section) may become attenuated or may be reversed in the bilateral version of the model to be considered in section 2 below.

2. Bilateral Accidents and Levels of Care

Now let us consider a bilateral version of the model of accidents, where victims as well as injurers can take care and thereby lower accident risks. The social goal will continue to be minimization of total social costs, which here will be the sum of injurers' as well as victims' costs of care, plus expected accident losses.

   2.1 Social welfare optimum. The optimal levels of care of injurers and of victims will reflect their joint possibilities for reducing accident risks and their costs of care. Consider the following example.

   Example 2. The probability of an accident that would cause losses of 100 is related to the different possible combinations of injurers' and of victims' levels of care as shown in Table 3. In this example, it is assumed for simplicity that there is only one positive level of care for parties of each type.

Table 3: Care of Injurers and of Victims, and Accident Risk						
Injurer care	Victim care	Injurer care cost	Victim care cost	Probability of accident	Expected losses	Total social costs
None	None	0	0	15%	15	15
None	Care	0	2	12%	12	14
Care	None	3	0	10%	10	13
Care	Care	3	2	6%	6	11

From the last column of the table, it is apparent that it is socially optimal for both injurers and victims to take care. To see why, observe, for instance, that if injurers alone take care, expected losses are 10, whereas if victims also take care, at a cost of 2, expected losses fall by 4; hence total social costs are reduced when victims also take care. Similar reasoning shows that the situation in which victims alone take care can be improved when injurers also take care.//

   Although in this example it is socially optimal for both injurers and victims to take care, other examples can obviously be constructed in which it is optimal only for injurers to take care or only for victims to take care (or for neither to do so). These possibilities are not the focus here (but see section 2.11) because in most real situations it would be best for both injurers and victims to take a positive degree of care, however small.[11]

---

[11]The formal version of the bilateral model is the natural extension of that of the unilateral model (see note 4): Injurers choose a level of care x, victims choose a level of care y, the probability of an accident is $p(x, y)$, which is declining in both x and y, and the social goal is to minimize $x + y + p(x, y)h$. It will generally be assumed, as just noted in the text, that the optimal levels of care x^* and y^* are positive and also that they are unique.

2.2 Behavior in the bilateral model. In the bilateral context, the way in which one type of party behaves will often depend on how the other type of party behaves. For example, how watchful drivers are for bicyclists may depend on how cautious bicyclists tend to be (drivers might be very watchful if bicyclists are not very cautious), and how cautious bicyclists generally are may depend on the usual attentiveness of drivers.

The possible interdependence of parties' actions means that if we want to show that some pattern of behavior will hold true, we have to show that it will be an *equilibrium* pattern in the sense that neither type of party would want to change what he is doing given the behavior of the other type of party. Injurers' and victims' behavior in equilibrium will now be determined in various liability settings.

2.3 No liability. As before, injurers will not take care in the absence of liability, and the outcome will therefore generally depart from the optimal. However, because victims bear their accident losses, they will have a reason to take care. In Example 2, although injurers will not take care, victims will take care, because for a cost of 2 they will lower their expected accident losses from 15 to l2. Note that this outcome is an equilibrium. It is in victims' interest to take care, given that injurers do not take care; and it is in injurers' interest not to take care, given that victims take care (or, for that matter, if they do not). The reader will be able to verify similarly that other outcomes below are equilibria, when this is not pointed out in the text.

2.4 Strict liability. Because injurers will be liable for the accident losses that they cause under strict liability, they will have a proper motive to take care. However, because victims will be fully compensated by injurers for accident losses, victims will be indifferent to the occurrence of accidents. Therefore, victims will not take care,[12] and the outcome will not be optimal. In Example 2, injurers will take care because doing so will reduce their expected liability from 15 to 10 at a cost of only 3, but victims will not take care.

2.5 Strict liability with the defense of contributory negligence. Under this rule an injurer is liable for the accident losses he causes only if the victim's level of care was at least equal to the victim's due care level. If the victim's care level was less than due care for him, the victim is said to be *contributorily negligent* and must bear his losses. (Contributory negligence is called a *defense* because it is a legal defense for the injurer: Its successful assertion by the injurer relieves him of liability.)

If courts choose the level of due care for victims to equal the socially optimal level of care, then victims will prefer to exercise due care and injurers also will prefer to take the socially optimal level of care. Thus, the socially optimal outcome will occur. To establish that this is true, note, first, that injurers will exercise optimal care, given that victims take due care, because then injurers will be liable for accident losses. (If bicyclists take due care, then drivers will be liable for accident losses and will decide to take optimal care.) Second, observe that victims will take

---

[12]However, victims would obviously have an incentive to take care if they would not or could not be compensated fully for their accident losses, as where the losses involve serious personal injury or death (which will be considered in section 8 of chapter 4 and section 6 of chapter 5). Thus, here (and often below) the reader may find it useful to think about examples of accidents in which victims would suffer only property losses. Nevertheless, the example in which victims are bicyclists will continue to be discussed in the text for expositional convenience. One might imagine, for example, that bicycle accidents damage bicycles but do not injure riders.

due care because they will wish to avoid being found contributorily negligent and thus having to bear their own losses. The specific reasoning is analogous to that in the explanation in section 1.4 of why injurers will take due care under the negligence rule.[13]

    To verify the claim in Example 2, assume that due care for victims equals "care," as victims' exercise of care is socially optimal. Presuming that victims take care, injurers will be liable for accident losses that they cause. Therefore their expected liability will fall from 12 to 6 if they spend 3 to take care, and they will take care. Conversely, assuming that injurers take care, victims will be induced to take care; for if victims do not take care, they will bear their expected accident losses of 10, whereas if they take care at a cost of 2 they will not bear their losses.[14]

    2.6 Negligence rule. As in the unilateral model, if the courts choose due care to equal the socially optimal level, then injurers will be led to take due care. Victims too will be induced to take the optimal level of care because they will bear their losses if injurers take due care. (Drivers will be led to take due care; and knowing that they will bear their losses, bicyclists will decide to take appropriate care.)

    To illustrate these conclusions, assume in Example 2 that due care for injurers equals "care." If injurers do not take care, their expected liability will be 12, presuming that victims take care; thus injurers will choose to avoid liability by spending 3 on care. Also, because victims will bear their losses when injurers take due care, victims will reduce their expected losses from 10 to 6 by taking care; as this will cost victims 2, they too will decide to take care.[15]

    2.7 Negligence rule with the defense of contributory negligence. According to this rule, an injurer will only be liable for accident losses if he failed to take due care and the victim exercised due care himself. In other words, if the injurer was negligent, he still will escape liability if the victim was contributorily negligent.

    An argument very close to that of the previous section shows that if courts choose

---

[13]This paragraph has explained only why both injurers and victims taking optimal care is *an* equilibrium. But the situation in which both take optimal care is in fact the *only* equilibrium that can exist. In other words, the only stable situation that can possibly exist under the rule of strict liability with the defense of contributory negligence (with due care for victims set at the optimal level) is that in which both injurers and victims take optimal care. That this equilibrium is unique follows from three observations: (1) Victims never have an incentive to take care y exceeding y^* (for once they take due care they will be compensated for their losses). (2) Victims will not choose y less than y^*, for if they do so, they will bear their own losses, injurers will take no care, and victims thus will minimize $y + p(0,y)h$. But $y + p(0,y)h = 0 + y + p(0,y)h > x^* + y^* + p(x^*,y^*)h > y^*$, implying that victims must be better off choosing due care y^* than any $y < y^*$. (3) Because in equilibrium victims thus take due care of y^*, injurers choose x to minimize $x + p(x,y^*)h$, which is minimized at x^*.

[14]To see why the only equilibrium in this example is the situation in which both injurers and victims take care, consider the other possibilities. For injurers to take care and for victims not to take care cannot be an equilibrium, since victims will wish to take care if injurers take care (or, also, if they do not). Similarly, for injurers not to take care and for victims to take care cannot be an equilibrium, since injurers will wish to take care given that victims take care. Finally, for both injurers and victims not to take care cannot be an equilibrium, since victims will wish to take care to avoid liability (for if they take care, their costs will be 2, whereas if they do not take care, they will bear expected losses of 15).

[15]The equilibrium in which both injurers and victims take optimal care is the only equilibrium under the negligence rule (assuming that due care is optimal). The socially optimal outcome is also the unique equilibrium under the next rule that we consider.

injurers' and victims' levels of due care to equal the socially optimal levels, both injurers and victims will be led to take due care and the socially optimal result will be achieved. Injurers will wish to take due care to avoid liability, under the assumption that victims take due care and thus will not bear their accident losses on account of contributory negligence. Also, victims will want to take due care, presuming that injurers take due care; since victims will then bear their losses, they will be led to take the socially optimal level of care, which by assumption is due care. (This may be verified in Example 2 exactly as it was in the preceding section.)

Notice that the defense of contributory negligence is a superfluous addition to the negligence rule with respect to the objective of inducing victims to act optimally, for it was seen in the last section that victims take optimal care when the negligence rule is unaccompanied by the defense. Under the negligence rule without the defense of contributory negligence, injurers take due care to avoid liability. Consequently, victims bear their losses, and this by itself supplies them an incentive to take appropriate care. Accordingly, there is no need to provide victims another incentive to take care.[16]

2.8 Comparative negligence rule. Under this rule, as under the last, an injurer will not be liable for accident losses he causes if he takes due care. But the comparative negligence rule differs from the previous rule in the situation in which *both* the injurer and the victim fail to take due care. In that case each party bears a fraction of the accident losses, where the fraction is determined by a comparison of the amounts by which the two parties' levels of care depart from the levels of due care. The fraction of losses a party bears will be higher the greater the difference between due care and his level of care.

If courts choose optimal levels of due care under the comparative negligence rule, then both injurers and victims will be led to take due care. The rationale for this conclusion is precisely that of the last section. (Injurers will take due care to avoid liability if victims take due care, and so on.)

The reason that there is no difference between the outcomes under the comparative negligence rule and under the negligence rule with (or without) the defense of contributory negligence is in essence this: Under both rules, if parties of one type take due care, then parties of the other type will reason that they alone will be found negligent if they fail to take due care. The allocation of accident losses when both injurers and victims are negligent -- the distinguishing feature of the comparative negligence rule -- therefore turns out to be irrelevant to the calculations of parties in equilibrium.[17]

2.9 Liability rules compared. We have seen that in the bilateral version of the model, strict liability does not lead to the socially optimal outcome for the obvious reason that it fails to furnish victims a motive to take care. We have also seen that strict liability with the defense of

---

[16]However, the defense of contributory negligence may generate beneficial incentives if some injurers act negligently. If some injurers act negligently and if there is no defense of contributory negligence, then a victim may decide not to take due care, since he may think he will be likely to obtain compensation for accident losses he suffers because they will be caused by a negligent injurer.

[17]But the allocation of losses when both injurers and victims are negligent *is* relevant in situations in which there are reasons why some injurers and victims act negligently; on such reasons, see sections 1 and 3 of chapter 4.

contributory negligence and all forms of the negligence rule result in the socially optimal outcome. Under these rules, parties have one of two sufficient reasons to take optimal care: Either taking optimal care allows them to avoid entirely the bearing of accident losses (victims' situation under strict liability with the defense of contributory negligence, injurers' situation under the negligence rules), or else taking care reduces the level of (rather than the entirety of) expected losses that parties in fact bear (injurers' situation under strict liability with the defense of contributory negligence, victims' situation under the negligence rules).

To apply each of the rules leading to optimality, courts need to determine the magnitude of accident losses and the actual level of care and the optimal level of due care for injurers or victims. Moreover, to ascertain the optimal level of due care for just one party, a court must generally determine (if only implicitly) the optimal level of care for the other as well, because the optimal level of care for one party will in principle depend on the other's costs of, and possibilities for, reducing risk.[18] This latter point makes the comparison of liability rules with respect to their ease of application different from what it might at first seem to be.

Consider, for instance, the rule of strict liability with the defense of contributory negligence and the negligence rule with the same defense. It may seem initially that strict liability with the defense of contributory negligence is the easier rule to apply, because courts are not directly concerned with injurers' behavior under the rule, whereas courts must set due care for injurers under the negligence rule. But to apply the defense of contributory negligence, courts must determine optimal due care for victims, and, as just remarked, this effectively requires courts to determine the optimal level of care for injurers. Therefore, the main difference affecting the ease of application of the two rules is only that under the strict liability rule courts do not need to observe the actual level of care of injurers.

2.10 Liability rules compared when care has several dimensions. I noted in section 1.6 that there may be dimensions of injurers' care (such as the frequency with which drivers look in their rear-view mirrors) that courts would not take into account in the determination of negligence because of difficulties in assessing them. Injurers may therefore not exercise care in an optimal way in every dimension under the negligence rule, but they will be led to do so under strict liability. It is clear that a similar point applies when there are dimensions of victims' care (such as the frequency with which bicyclists look for traffic behind them) that could not be included in their standard of due care. Specifically, victims will not take optimal care in these dimensions under strict liability with the defense of contributory negligence, but they will do so under the negligence rule (because they will bear their accident losses under that rule). In consequence, to know how the presence of multiple dimensions of care affects the comparison of liability rules, one must make a judgment about the relative importance of the dimensions of injurers' and of victims' behavior that would be excluded from their respective standards of due

---

[18]That courts must generally consider the entire tableau of costs and effectiveness of care for the two parties to determine optimal care for either should have been evident from Table 3 and section 2.1. But it should be mentioned that in some situations the optimal level of care for parties of one type may be determinable without precise knowledge of the other's optimal level of care. Suppose, for instance, that the use of lights by bicyclists when riding at night will dramatically reduce accident risks whatever the level of care taken by drivers. Then it would be optimal for bicyclists to use lights at night without determining what particular level of care is optimal for drivers.

care.

2.11 The least-cost avoider. The notion of the *least-cost avoider* applies in situations in which the risk of accidents will be eliminated if *either* injurers or victims take care. In such situations it is clearly wasteful for *both* injurers and victims to take care; rather, it is optimal for the type of parties who can prevent accidents at least cost -- the least-cost avoiders -- alone to take care. Suppose, for example, that injurers can prevent accident losses of 100 by taking a precaution that costs 10, and that victims also can prevent the losses by taking a precaution that costs 20. In this case injurers alone ought to take precautions, because in that way the social goal of minimizing total social costs is achieved.

The model of the least-cost avoider may be misleading for thinking about the class of bilateral accidents examined in this book. In the situations examined here, there simply are no least-cost avoiders who alone ought to take care, for the assumption is that both injurers and victims generally ought to do something to avoid risk; the effect of liability rules is therefore different from that in the least-cost avoider model. If, say, injurers are the least-cost avoiders, an optimal outcome will be achieved under strict liability unaccompanied by the defense of contributory negligence. But in the bilateral model studied here, the defense of contributory negligence must accompany strict liability in order to induce victims as well as injurers to take appropriate care.

2.12 Liability rules in use. The major rules of liability for accidents between strangers in the United States are the comparative negligence rule, the negligence rule with the defense of contributory negligence, and strict liability with that defense.[19] In England, France, and Germany, the usual forms of liability are the comparative negligence rule and strict liability with forms of contributory negligence defense.[20]

2.13 The determination of due care and the *as if* interpretation. Negligence in American law, according to the *Restatement (Second) of Torts* is "conduct which falls below the standard [of due care] . . . for the protection of others against unreasonable risk of harm," and the concept of negligence is similar in other legal systems. Deciding on the standard of due care often requires some sort of weighing of the magnitude of risk against the disutility or cost of more careful conduct.[21]

As the reader has seen in the analysis here, the level of due care that minimizes total social costs implicitly involves just such a weighing of risk against the cost of care. This suggests that due care is in fact found by a process that operates as if it were designed to identify

---

[19]See, for example, Keeton, Dobbs, et al. 1984, chapters 5, 11, and 13; and Dobbs 2000, chapters 6, 11, and 23.

[20]See Tunc 1983, for a summary of and bibliography on tort law in the entire world; Fleming 1998, for a treatment of tort law focusing on England and Australia; Von Bar, 1998, vols. 1 and 2, for a description of tort law in Europe; Von Mehren and Gordley 1977, chapters 8-10, for materials on tort law in France and Germany; and Zweigert and Kötz 1998, chapters 40-43, for a description of tort law in England, France, and Germany.

[21]See the *Restatement (Second) of Torts* 1965, sections 282, 291-93. The *Restatement* is a summary of and commentary on the doctrines of tort law produced by leading scholars under the aegis of the American Law Institute. For discussion of the determination of negligence in other legal systems see, for example, Limpens et al. 1983, sections 23-27; Markesinis 1994, 72-74; Von Bar 1998, 1: 20-39, 2: part 2; and Zweigert and Kötz 1998, 599-600, 615-17.

behavior that minimizes total social costs, or at least approximately so (one does not know if the weighing is any more than qualitatively similar to that which minimizes total social costs).[22]

I use the words "as if" because the claim is hardly that individuals or courts *think* in terms of the mathematical goal of minimizing a sum. They obviously do not do anything so unnatural. Rather, they appear to gauge the appropriateness of behavior by a rough consideration of risk and the costs of reducing it, ordinarily on the basis of felt notions of fairness.[23] Likewise, the *as if* interpretation carries with it no specific implications about the degree to which individuals and courts concern themselves about goals of deterrence, although they sometimes appear to be concerned about it.

With these caveats in mind, observe that the *as if* interpretation is borne out not only by the mere fact that there is a weighing involved in the negligence determination, but also by a consideration of the character of the weighing. First, the list of factors that courts take into account in setting due care -- and the influence of those factors on the level of due care -- are what we would expect if courts were aiming to minimize total social costs: The level of due care is generally higher the greater the likelihood of harm, the larger the probable size of harm, the greater the number of individuals at risk, and the easier it is for injurers to alleviate risk.[24] Second, the choice of due care levels probably reflects the possibilities for both injurers and victims to reduce accident risks, as is consistent with the bilateral model of accidents. Consider, for instance, the risk of accidents in which bicyclists run into car doors as the doors are opened. My surmise is that most of us would say that bicyclists should not have to proceed so slowly that, were a car door to open suddenly, they could virtually always stop in time; rather, we would say that, before persons open their car doors, they should look around to see if anyone is approaching. I suggest too that in coming to this view, most of us would have at the back of our minds -- if not in our conscious thoughts -- such ideas as that it would be a burden for bicyclists to have to go so slowly that they could stop immediately before running into car doors, that it is relatively easy for persons leaving cars to look for danger, and that it is not necessary for bicyclists to go very slowly if persons are properly cautious when leaving their cars. In other words, when deciding on the care that parties of one type ought to exercise, we quite naturally factor into our thinking the ability of parties of the other type to take care, the burden of doing so, and what their taking care would accomplish.

Note on the literature. The first writer to study in an analytical way the theory of the

---

[22]It will be clear to the reader that the "as if" interpretation can also be made about many other instances that will be mentioned in later chapters of this book of consistency between actual law and the law that is theoretically optimal (given the stated measure of social welfare).

[23]An exception is Judge Learned Hand's algebraic formula for determining the due care standard. In his opinion in *United States v. Carroll Towing Co.*, 159 F.2d 169 (2d Cir. 1947), Hand said that a party is negligent if he failed to take a precaution when its cost, which he called its "burden," was less than its expected benefit; he denoted the burden by B, the probability of loss in the absence of precaution by P, the magnitude of loss by L, and said that negligence should be found if $B < PL$.

[24]See Fleming 1998, chapter 7, section 2; Keeton, Dobbs, et al. 1984, sections 29, 31, 33; the *Restatement (Second) of Torts*, 1965, section 293. Note too that these effects on due care are consistent with Judge Hand's formula.

effect of liability rules on parties' behavior was Calabresi (1961, 1965, 1970). He examined the desirability of different rules, emphasizing the advantages of versions of strict liability, and assuming for the most part the goal of minimization of total social costs.[25] Posner (1972a,b, l973) later made significant contributions, especially in his analysis of the various principles and doctrines governing use of the negligence rule.[26] Although both these writers discussed suggestive numerical examples, neither recognized that liability rules would, as a general matter, lead calculating parties to choose levels of care such that total social costs are minimized. Brown (1973) put forward the first clear statement and formal proofs of this conclusion. He showed that the rules of strict liability with the defense of contributory negligence and the negligence rule (with or without the defense) induce injurers and victims to take optimal levels of care in equilibrium.[27]

3. Unilateral Accidents: Levels of Care and Levels of Activity

We will now consider an injurer's *level of activity* -- that is, whether, or how much, he engages in a particular activity. The number of miles an individual drives, for instance, might be interpreted as his level of activity. An injurer's level of activity is to be distinguished from his level of care, which has to do with the precautions he takes *when* engaging in his activity (the precautions an individual takes when on the road, such as slowing for curves, as opposed to the number of miles he drives).

Our analysis will begin with the unilateral case, and we will assume for simplicity that an increase in an injurer's activity level will result in a proportionate increase in expected accident losses, given his level of care. Thus, a doubling in the number of miles that individuals drive will result in a doubling in the number of accidents they cause, given the care with which they drive; or a doubling in the number of times individuals walk their dogs will result in a doubling in the risk that their dogs will bite strangers, given the care (leashing the dogs) they take to prevent attacks. We will also assume that an increase in an injurer's level of activity will result in an increase in his utility (at least up to some point); the more individuals drive or the more they walk their dogs, the greater will be their utility (until their need to drive is met or until walking their dogs turns into a chore).

We will now assume the social goal to be maximization of the utility injurers derive from engaging in their activity less total social costs, that is, less the costs of care and expected accident losses. It makes sense, of course, to introduce the utility injurers derive from their

---

[25]Many previous writers had, of course, recognized that liability rules would have some effect on behavior, but usually only in passing. Calabresi differed from his predecessors in that he made the effect of liability on behavior the focus of his work and carried it out in a self-conscious, sustained, and careful way.

[26]See also Landes and Posner 1981a, 892-903, discussing what is called here the *as if* interpretation of the negligence determination.

[27]Soon afterward Diamond 1974a,b also showed, in closely related models, that the negligence rule with the defense of contributory negligence induces parties to take levels of care that minimize total accident cost. See as well Green 1976, Shavell 1987a, 72-77, 86-91, Emons 1990, and Emons and Sobel 1991, who analyze liability when injurers and victims are heterogeneous.

activity into the measure of social welfare, because the level of their activity is now a subject of study.[28]

3.1 Social welfare optimum. For social welfare to be maximized, an injurer must, as before, choose a level of care that is commensurate with the effect of care in reducing accident losses and with its costs. But now the injurer should also select his level of activity appropriately, which is to say, at the level that appropriately balances the utility he obtains against the additional risks he creates and the costs of care.

Example 3. Let us build on Example 1 by assuming that it describes the situation each time injurers engage in their activity. Thus, injurers who behave optimally will take moderate care, at a cost of 3, and will reduce expected accident losses to 10. Consequently, if an injurer engages in his activity twice, taking optimal care each time, his total costs of care will be 6, and the expected accident losses he causes will be 20; if he engages in his activity three times, the figures will be 9 and 30, respectively; and so forth. These figures are shown in the third and fourth columns of Table 4. The second column in the table shows the total utility injurers derive from engaging in the activity, from which the figures for social welfare in the last column -- utility minus costs of care and accident losses -- can be calculated.

Table 4: Activity Level, Accidents, and Social Welfare				
Activity level	Total utility	Total costs of care	Total accident losses	Social welfare
0	0	0	0	0
1	40	3	10	27
2	60	6	20	34
3	69	9	30	30
4	71	12	40	19
5	70	15	50	5

The optimal activity level is 2 because social welfare is highest at that level. One way of explaining why is as follows. Each time an injurer engages in the activity, he will increase total social costs by $3 + 10 = 13$. Therefore, social welfare will be enhanced by his engaging in the activity another time if and only if the marginal utility he would gain exceeds 13. Because the utility he obtains from engaging the first time is 40, the marginal utility he obtains from the second time is 20 (that is, $60 - 40$), and that from the third time is only 9 (that is, $69 - 60$), it is

---

[28]The social goal considered above, minimizing total social costs, may be viewed as a special case of the present goal. If we imagine the level of activity, and hence the utility from the activity, to be held constant, as we implicitly assumed was the case above, then maximization of the utility derived from the activity less total social costs is obviously equivalent to minimization of total social costs.

best that he stop at the second time.[29] //

The general point illustrated by this example is that the socially optimal behavior of injurers can be determined in two steps: first by finding (as in section 1.1) the level of care that minimizes total social costs incurred each time injurers engage in their activity, and then by raising the level of activity as long as the marginal utility injurers derive exceeds the increment to total social costs.[30]

3.2 No liability. In the absence of liability, not only will injurers fail to take care, they also will engage in their activity to too great an extent. Indeed, they will continue to engage in it as long as they obtain any additional utility (individuals will go for a drive or walk their dogs on a mere whim) rather than, as would be socially desirable, only as long as they obtain additional utility exceeding the costs of optimal care plus the expected accident losses they cause. In Example 3 injurers will not take care and thus will choose activity level 4, the level at which they cease to gain utility from their activity, rather than the optimal activity level of 2.

3.3 Strict liability. Under strict liability an injurer's utility, net of his expected costs, will be equal to the measure of social welfare, because he will pay for the accident losses he causes, will naturally enjoy the benefits of engaging in his activity, and will bear the costs of care. Accordingly, injurers will behave so as to maximize social welfare; they will thus choose both the optimal level of care and the optimal level of activity.

More directly, injurers will choose the optimal level of care because doing so will minimize the expected costs they bear each time they engage in their activity. And they will choose the optimal level of activity because they will wish to engage in the activity only when the extra utility they derive exceeds their costs of care plus their added expected liability payments for accident losses caused. (People will walk their dogs only when their utility gain outweighs the disutility of having to leash the dogs and the added liability risk due to dog bites.) In Example 3, we know (from section 1.3) that strictly liable injurers will take the moderate level of care. Hence, the last column in Table 4 will become injurers' utility, net of their expected liability costs, and they will therefore choose the optimal activity level of 2.

3.4 Negligence rule. As the reader recalls from previous analysis, injurers will be led to take optimal care under the negligence rule, assuming that courts choose the level of due care to equal the optimal level of care. Because they will take due care, however, injurers will escape liability for any accident losses they cause. They will therefore have no reason to consider the effect that engaging in their activity has on accident losses.

Consequently, injurers will be led to choose socially excessive activity levels.

---

[29]Notice that utility actually falls beyond activity level 4. (The fifth time one walks his dog, it is more a chore than a pleasure.)

[30]The formal model illustrated in the example is as follows. Let z be the activity level, $b(z)$ be the utility or benefit from the activity, and assume the social object is to maximize $b(z) - z(x + p(x)h)$, where $x + p(x)h$ are social costs each time an injurer engages in his activity. Let z^* and x^* be optimal values of z and x. Note that x^* minimizes $x + p(x)h$, so x^* is as described above in section 1; x^* minimizes social costs whatever the level of activity. Therefore, z^* is determined by $b'(z) = x^* + p(x^*)h$, which is to say, the condition that 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).

Specifically, they will engage in their activity whenever the utility they derive net of the cost of care is positive (whenever the pleasure from walking their dogs net of the disutility of leashing them is positive), rather than only when their net utility exceeds the additional expected accident losses they create.[31]

This can be seen in Example 3, where we know that if due care is the optimal, moderate level, injurers will take due care. Because injurers take due care under the negligence rule, they will not be liable for accident losses and their situation will be that described in Table 5.

Table 5: Negligence Rule and Activity Level			
Activity level	Total utility	Total costs of care	Total utility – costs of care
0	0	0	0
1	40	3	37
2	60	6	54
3	69	9	60
4	71	12	59
5	70	15	55

From the last column in the table it is evident that injurers will choose the activity level 3 rather than the optimal activity level 2: They will increase their activity level from 2 to 3 because this will raise their utility by 9 and their costs of care by only 3; they will not consider that increasing their activity level will also raise expected accident losses by l0 (as shown in Table 4), for they will not be liable for these additional social costs.

3.5 Liability rules compared. Under both strict liability and the negligence rule, injurers are led to take socially optimal levels of care, but under the negligence rule, they engage in their activity excessively because, unlike under strict liability, they do not pay for the accident losses that they cause.

The importance of this defect of the negligence rule will clearly depend on the expected magnitude of the losses caused by an activity. If an activity is by its nature very dangerous even when carried out with appropriate precautions, it will be significant that under the negligence rule the level of the activity would be socially excessive. For example, if the walking of dogs of a vicious breed or if blasting creates high risks of harm despite the use of all reasonable (or due) care, it will be of real consequence that under the negligence rule people would walk their dogs

---

[31]In terms of the model (see the previous note), the point is as follows. Because an injurer will escape liability by exercising care x^*, he will choose z to maximize $b(z) - zx^*$, so that z will satisfy $b'(z) = x^*$. However z^* is determined by $b'(z) = x^* + p(x^*)h$, so that z will be excessive under the negligence rule (assuming that $b(z)$ is a concave function).

excessively (rather than exercise them in a yard or rather than own dogs of another breed) or that firms would blast excessively (rather than employ other methods of demolition). If, however, an activity creates only a low risk of accidents when due care is taken, the importance of any excess in the level of activity under the negligence rule will be small. This is true, one suspects, of many, and perhaps most, of our everyday activities (mowing a lawn, playing catch, walking the friendly, domesticated dog).

 3.6 The source of the defect of the negligence rule. The failing of the negligence rule results from an implicit assumption that the standard of behavior for determining negligence is defined only in terms of the level of care,[32] an assumption that seems generally to be true in reality. Were the negligence standard defined so as to include the activity level, injurers would make sure not to engage in their activity to an excessive extent in order to avoid a finding of negligence.

 This consideration, however, immediately raises the question why the courts do not usually include the activity level in the determination of negligence. A possible answer concerns the information that the courts would require. To formulate a standard for the level of activity, courts would need to ascertain the character of the benefits that parties derive from their activities. (Courts would have to inquire into the pleasure obtained from walking a dog or the importance of driving somewhere.) Because these benefits often seem practically unknowable, attempts by courts to determine appropriate levels of activity would probably land them in a speculative realm. Deciding on appropriate levels of care, although by no means an easy task, usually appears to be less problematic. (We can say with fair confidence that a dog that tends to snap at others should be leashed, or that a person should not drive at 60 miles per hour along a residential street.)

 Aside from the difficulties that courts would face in formulating appropriate standards for parties' levels of activity, courts would have to verify what parties' levels of activity actually were. This additional burden might be a substantial one in some situations, especially because establishing an individual's level of activity would require knowledge of what he did in the past. (How many times did a person walk his dog before the last time, when it bit someone?) By contrast, assessing an individual's level of care often requires knowledge of his behavior only at the time of an accident.

 Nevertheless, there may be situations in which a court would have sufficient information to incorporate the level of activity into the negligence determination. One notable example is when a party engages even once in an activity that is very dangerous despite the exercise of care, and the activity yields the party only an obviously small utility. In this case the party could be called negligent merely for having engaged in the activity.[33]

---

[32]Notice therefore that the defect is similar to that discussed in section 1.6 concerning dimensions of care omitted from the due care standard.

[33]In this regard, it is interesting to note the passage in the *Restatement (Second) of Torts* 1965, section 297, which reads in part, "A negligent act may be one which involves an unreasonable risk of harm . . . although it is done with all possible care." By way of example, the *Restatement* comments that "there are many mountain roads which may properly be regarded as dangerous no matter how careful . . . the driver may be . . . there is an inescapable risk in driving down a narrow and illkept mountain road . . . particularly if . . . snow or ice has rendered the road slippery . . . mere use of such a

Note on the literature. In Shavell (1980c) I introduced the distinction between the level of activity and the level of care and first developed the points about strict liability versus negligence and the activity level that are discussed here and in section 4 below.

4. Bilateral Accidents: Levels of Care and Levels of Activity

In this most general case victims as well as injurers will be assumed to choose levels of activity and levels of care. As with injurers' levels of activity, increases in victims' levels of activity will be assumed to raise their utility, at least up to some point, and will result in proportionate increases in expected accident losses. Thus, if a bicyclist rides an extra mile, he will enjoy extra utility and his chances of being involved in an accident will rise. The measure of social welfare will be taken to be the utility that victims and injurers derive from their activities less their costs of care and expected accident losses.

The analysis that follows will be brief because most conclusions can be explained by appeal to the previous cases.

4.1 Social welfare optimum. Optimal behavior in the bilateral case will reflect not only the cost of care and its effect on accident risks, but also the utility that injurers and victims obtain from their activities.

Example 4. Suppose for simplicity that victims either engage in their activity or they do not, and suppose the same for injurers; in other words, for parties of each type, there is only one possible positive level of activity. Suppose also that if parties of one type engage in their activity and the others do not, no accidents can occur -- it takes the presence of both injurers and victims for there to be accidents. Hence, if parties of only one type engage in their activity, it would be pointless and socially wasteful for them to take care. Finally, suppose that if both injurers and victims engage in their activities, the risk of accidents will be as described in Example 2. Thus in this case injurers ought to take care, which costs 3; victims also ought to take care, which costs 2; and expected accident losses will be 6. Therefore, total social costs will be $3 + 2 + 6 = 11$ if both injurers and victims engage in their activities and take care.

Given these assumptions, it is easy to determine when it is optimal both for injurers and for victims to engage in their activities, as a function of the utilities they would each derive from so doing. Were parties of only one type to engage in their activity, none of the accident costs of 11 would be borne (because no accidents could occur and no care would be taken). Therefore, it will maximize social welfare for both injurers and victims to engage in their activities only when each would obtain from their activity a utility exceeding 11. Otherwise, it will be best for the parties that would enjoy the greater utility to engage in their activity and for the other parties to refrain from engaging in their activity.

To verify this claim, suppose for instance that injurers would obtain utility of 35 and victims 25 from engaging in their activities. If both injurers and victims engage in their activities, social welfare will be $35 + 25 - 11 = 49$; if only injurers engage in their activity, social welfare will be 35; if only victims do so, social welfare will be 25; thus it will indeed be optimal for both injurers and victims to engage in their activities. Next suppose that injurers would

route . . . may be negligent unless the utility of the route is very great."

obtain 35 from engaging in their activity and victims would obtain only 8. Then if both injurers and victims engage in their activities, social welfare will be $35 + 8 - 11 = 32$; if injurers alone do so, social welfare will be 35; if victims alone do so, social welfare will be 8; and it will be best for injurers alone to engage in their activity.[34] Similar calculations show that if injurers would obtain 8 and victims 25 from engaging in their activities, then it will be optimal for victims alone to engage in their activity.//

The simplifying feature of this example, that parties either do not engage in their activity or engage in it at only one positive level, should not disturb the reader. The points to be illustrated below will carry over in obvious ways to the more realistic case in which there are many different positive levels of activity for each type of party.

4.2 Strict liability with the defense of contributory negligence. As the reader knows from previous analysis, if courts select the optimal level of due care, then under strict liability with the defense of contributory negligence, both injurers and victims will be led to take optimal care when they engage in their activities. Furthermore, because victims will take due care, injurers will pay for the accident losses they cause and thus, as is explained in section 3.3, they will choose the correct level of their activity given victims' behavior.

Yet because victims will be compensated for their losses, victims may engage in their activity too often. A victim's only cost of engaging in his activity will be his cost of taking due care. Therefore, he will engage in his activity whenever his utility from so doing would exceed the cost of taking due care. But what would be desirable is that he engage in his activity only when his utility would exceed the cost of taking due care plus the expected accident losses that would result from his engaging in his activity. (A bicyclist will go for a ride whenever the pleasure he would gain from that ride exceeds the disutility from having to exercise appropriate care, rather than only when the pleasure exceeds the disutility of exercising such care plus the increment to expected accident losses.)

To illustrate this point, consider the case in Example 4 in which injurers would obtain utility of 35 and victims utility of only 8 from their activities, and thus in which it is not optimal for victims to engage in their activity. Under strict liability with the defense of contributory negligence, victims need only take due care, at a cost of 2, to be assured of compensation for accident losses suffered. Hence, when they compare the utility of 8 that they would obtain from engaging in their activity to the cost of care of 2, victims will, undesirably, decide to engage in their activity (along with the injurers, who will compare their utility of 35 to their cost of care of 3 plus their expected liability of 6).

4.3 Negligence rule with or without the defense of contributory negligence. Again, the reader knows from previous discussion that under the negligence rule, both injurers and victims will be induced to take optimal care when engaging in their activities if courts select optimal due care levels. And since injurers will escape liability by taking due care, it is evident from the argument of section 3.4 that injurers may engage in too high a level of their activity.

---

[34]It is not necessary that injurers enjoy utility greater than 11 for it to be optimal for them to engage in their activity. For instance, if injurers' utility were 10, it would still be optimal for them alone to engage in their activity, and social welfare would be 10.

Victims, however, will choose the correct level of their activity given injurers' behavior. Because victims will bear their own losses, they will engage in their activity a further time only if the utility they would obtain (net of the costs of taking care) exceeds the addition to expected losses. Consider the situation in Example 4 in which injurers would obtain utility of 8 and victims utility of 25 from engaging in their activities. In this case it is optimal for victims alone to engage in their activity, and under the negligence rule they will do so (for they will compare 25 to 2 + 6), but so will injurers, undesirably (for they will compare 8 to 3).

4.4 Liability rules compared. It should be evident from what has been said that strict liability with the defense of contributory negligence will result in higher social welfare if its disadvantage -- that victims engage too often in their activity -- is not as important as the disadvantage of the negligence rules -- that injurers engage too often in their activity. That is, strict liability will result in greater social welfare if it is more important for society to control injurers' levels of activity than victims'.

Whether injurers' levels of activity are more important to control than victims' will depend on the context. As discussed before, when an activity of injurers (walking dogs of a vicious breed) creates substantial risks despite their exercise of due care, the activity will be desirable to control. This point is not fundamentally altered if account is taken of the activities of victims that expose them to risk. Especially if the victims' activities are just the activities of ordinary life (walking about, going to work), we would not want the activities constrained in favor of injurers' more dangerous activities. Conversely, when an activity of injurers (playing baseball) is not very dangerous if appropriate care is taken, the importance of controlling the activity will not be great; instead, we may see some advantage in reducing certain activities of victims that subject them to particular risks (such as pushing a baby in a stroller across a baseball field while a game is in progress).

4.5 Nonexistence of a liability rule leading to optimal levels of activity. Because neither of the liability rules, strict liability and negligence, induce both injurers and victims to choose optimal levels of their activities, one might ask whether there exists any conceivable liability rule that always results in optimal levels of activities. The answer is no. The reason, in essence, is that for injurers to choose the correct level of their activity they must bear accident losses, whereas for victims to choose the correct level of their activity they too must bear accident losses. Yet it is not possible for both injurers and victims to bear accident losses under a liability rule.[35]

Three comments should be made about this conclusion. First, the explanation just given for it directly suggests methods (different from liability rules) that in principle would lead to optimal behavior. For example, suppose that injurers pay fines to the state equal to harm done -- or taxes equal to expected harm -- and that victims bear their losses. Then the expected payments of injurers and of victims would each equal expected accident losses, and they would each choose optimal levels of their activity (as well as care). Second, the conclusion depends on the

---

[35]The specific conclusion described in this paragraph is as follows. Assume that a liability rule can depend only on the levels of care x and y of parties and harm h, but not on their levels of activity. Then the rule cannot in general induce both injurers and victims to choose optimal levels of activity; see Shavell 1980c.

assumption that courts cannot incorporate parties' levels of activity into the negligence or the contributory negligence determination (an assumption that may be justified by what was said in section 3.6). If negligence and contributory negligence could be defined in terms of levels of activity as well as levels of care, then the usual liability rules would lead injurers and victims to choose optimal levels of both care and activity. Third, the conclusion should not be interpreted as an unduly negative one. As more factors are incorporated into a model, it naturally becomes less likely that a hypothetically ideal outcome can be achieved.

4.6 The reciprocal nature of harm. It is a truism that harm has a reciprocal aspect in the sense that a victim must be present to suffer harm just as much as an injurer must be present to do harm. This observation has sometimes been said to imply that injurers should not necessarily pay for harm done, that harm should not necessarily be "internalized" to injurers. That conclusion is supported by the analysis here, for as explained, either strict liability or negligence rules could turn out to be best.

The reciprocal nature of harm has also occasionally been suggested to mean that it is conceptually impossible to decide whether strict liability or the negligence rule should be applied, and even that the very notion of harm and its cause is rendered ambiguous.[36] This view is mistaken. There is no difficulty in principle in deciding whether strict liability or the negligence rule will be better in a given situation in a well-defined model (there was no difficulty in deciding the question with regard to Example 4, for instance),[37] and there is nothing problematic about the notion of harm.

4.7 Actual use of strict liability and negligence rules. The choice between the two main forms of liability for accidents between strangers has been made in approximately the same manner in different legal systems.[38] Namely, negligence is the usual basis of liability; strict liability applies only in certain areas of accident. In Anglo-American law, liability for accident losses is "for most significant purposes governed by the concept of negligence"; use of strict liability is restricted to harms caused by wild animals, to certain types of harms due to fire, and to harms arising from "abnormally dangerous" or "ultrahazardous" activities, such as blasting, storage of flammable liquids, or transport of nuclear materials.[39] Most of the provisions of the German civil code impose liability only if the injuring party was at fault; strict liability is adopted in connection with harms due to animals other than domestic animals and, according to special legislation, in connection with harms arising from rail, road, and air traffic and from use

---

[36]See Donohue 1989, 1057.

[37]But that there are no difficulties in principle does not mean that there will be no difficulties in application. Suppose that just as a woodsman cuts down a tree, a hiker happens to come along and is struck by the tree. Here we might feel that there is no appealing notion of who ought to have been present because it is hard to make a relative judgment about the benefits the hiker and the woodsman derived from their activities.

[38]Liability for accidents involving firms and their customers will be discussed in the next chapter.

[39]See Dobbs 2000, chaps. 6, 23; Fleming 1998, chapters 6, 15-18, quotation p. 97; and Keeton, Dobbs, et al. 1984, chapters 5, 13.

of electricity, gas, and atomic energy.[40] The situation in France is similar.[41] Two important articles of the French civil code specify fault or negligence as the general principle of liability; strict liability applies to harms due to animals or to certain dangerous things (including automobiles and aircraft).

4.8 Strict liability and negligence rules in the light of the theory concerning levels of activity. As stressed in the analysis, the use of strict liability rather than negligence rules in areas of behavior where activities create high risks, despite the exercise of reasonable care, has the advantage that it will tend to reduce in a desirable way participation in these activities.

This theoretical advantage seems consistent with reality in the sense that the impression given by the foregoing section is that the areas of activity covered by strict liability are *generally* more dangerous than those covered by negligence rules (certainly the reverse is not true). There are some exceptions to this pattern, however; the choices made between strict liability and negligence rules are not always easy to explain on the basis of differences in riskiness. (In the United States, is the danger due to the escape of wild animals from zoos, for which strict liability would probably apply, greater than that from automobile-pedestrian accidents, for which the negligence rule would govern?) Moreover, differences among countries in the areas of strict liability and of negligence are sometimes difficult to explain in terms of differences in dangerousness. (Why should the negligence rule govern liability for automobile-pedestrian accidents in the United States, while strict liability applies in Germany and France?)[42] The conformity of the observed pattern of use of strict liability and negligence rules to what would be suggested by the theoretical considerations of this chapter is somewhat rough.

Putting aside questions concerning the actual dangerousness of the areas of strict liability versus those of negligence, I want to emphasize that one of the aims of the law is to impose strict liability on activities that are dangerous, or, more precisely, that are dangerous even if conducted with reasonable care. A particularly direct expression of this objective is provided by the *Restatement* (*Second*) *of Torts*, which says that, in deciding whether an activity should be subject to strict liability, one ought to take notice of possible "inability to eliminate the risk by the exercise of reasonable care." Further, the *Restatement* draws a contrast to most "ordinary activities" that can be made "safe by the taking of all reasonable precautions" and for which liability should be based on negligence.[43]

But it should be added that the deterrent effect of strict liability on the level of participation in activities is not mentioned in the *Restatement* and is only infrequently noted in other places. Evidently, the mere creation of an unusual risk is seen as a justification for

---

[40]See Limpens et al. 1983, sections 11-14; Markesinis, 1994, 676-720; Opoku 1972, 230-43; Tunc 1983, sections 13, 79-85; Von Mehren and Gordley 1977, 557-66, 579-82; and Zweigert and Kötz 1998, chapter 42.

[41]See Limpens et al. 1983, sections 5, 23; Tunc 1983, sections 12, 86-88; Von Mehren and Gordley 1977, 555-57, 579-82; and Zweigert and Kötz 1998, chapter 42.

[42]On Germany, see Markesinis 1994, 710-20; and Opoku 1972, 240; on France, see Von Bar, 1998, 2: 410 – 11.

[43]*Restatement* (*Second*) *of Torts* 1965, section 520.

imposition of strict liability.

4.9 Note on empirical literature. A somewhat limited amount of empirical work has been undertaken on the effect of liability on accidents. See Dewees, et al (1996) for a general survey and, among others, Devlin (1990), Landes (1982), Sloan, et al (1994), Sloan (1998), and Cummins, et al (2001) on automobile accidents.

Chapter 3

LIABILITY OF FIRMS

In this chapter I reconsider the basic theory of liability and deterrence under the assumption that injurers are firms.[44] The first part of the analysis addresses accidents in which the victims are strangers to firms, such as an accident in which a gasoline tanker truck crashes and explodes, harming other vehicles or homes near the roadside.

The second part of the analysis deals with accidents in which the victims are the customers of firms, for example, an accident in which a water heater that a person purchased ruptures and damages his property. The feature of chief interest about these situations is that customers' willingness to purchase products will be influenced by what they perceive to be the product risks. As a consequence, firms will be motivated to reduce product risks not only to avoid liability but also to sell products at a better price.

Firms will be presumed to maximize profits and to do business in a perfectly competitive environment. This means that the price of a product will equal the unit costs associated with production, including expected liability costs.[45]

The measure of social welfare that will be studied is similar to that in chapter 2: the utility customers derive from products (such as from gasoline or water heaters) and, where relevant, the utility that strangers obtain from their activities (such as driving, or locating their homes near roadsides), minus expected accident losses, the costs of care, and direct costs of production.[46]

---

[44]The material here is based largely on chapter 3 of Shavell 1987a, which also contains proofs of conclusions.

[45]Under perfect competition, a firm cannot maintain a product price exceeding total unit costs because competitors could then attract its customers by offering them the same product at a lower, yet still profitable price. The assumption of perfect competition is made mainly for convenience -- it means that the effect of changes in unit costs on price is simple to calculate. See section 2.9 for further discussion of this assumption.

[46]This measure of social welfare should seem, on its face, a natural one to study, as it takes into account the obvious social benefits and costs associated with production of a risky product. Additionally, the measure is equivalent to the sum of the utilities of all relevant parties. In particular, the measure is the sum of the utilities of customers, owners of firms, and strangers in the case in which strangers are accident victims (the measure is the sum of the utilities of just the customers and the owners of firms in the case in which customers are accident victims). To verify this, observe that: Utility of customers = utility from product – price; utility of owners of firms is firms' profits, and firms' profits = price – direct production costs – costs of care – expected liability payments; utility of strangers = utility from their activity – costs of care – expected accident losses + expected liability payments. Adding these equalities and canceling offsetting terms, one obtains: Utility of customers + profits of owners of firms + utility of strangers = customers' utility from product – direct production costs – firms' cost of care + strangers' utility from their activity – strangers' costs of care – expected accident losses, which is as claimed.

1. Victims Are Strangers to Firms

Although the conclusions to be drawn about liability rules will be essentially the same in this case as in the last chapter, there are several differences in how the conclusions are demonstrated and interpreted that merit attention.

1.1 Levels of care. The arguments given in chapter 2 with respect to parties' levels of care apply directly in the present case. Victims are in an identical situation whether injurers are firms or are other individuals; and, as before, injurers, who in this case are firms, want to minimize their costs of care plus expected liability expenses (by doing so, firms maximize their profits). Hence, both firms and victims will be led to take optimal levels of care under strict liability with the defense of contributory negligence and under the various negligence rules.

The type of liability rule employed will, however, affect product price, a factor that was not relevant in the last chapter. In particular, the price will be higher under the strict liability rule than under the negligence rules. Under strict liability, the price will include expected accident losses, whereas under the negligence rules, it will not because, by taking due care, firms will avoid liability for accident losses.[47]

Example 1. Firms' direct costs of production per unit are 10, and the risk of accidents that would cause losses of 100 depends on whether firms take care (for simplicity, we consider the case in which accidents are unilateral). The exercise of care reduces expected accident losses by 6 and raises costs by only 2, as shown in Table 1. Thus, it is socially desirable for firms to take care.

Table 1: Care of Firms and Accidents				
Care Level	Cost of care	Probability of accident	Expected accident losses	Total social costs
None	0	9%	9	9
Care	2	3%	3	5

Under the negligence rule, firms will have to take care to avoid liability. Firms therefore will take care, and their costs per unit will be 12 -- the direct production costs of 10 plus the costs of care. Accordingly, the product price will also be 12 (by assumption, competition will drive the price down to unit costs).

Firms will take care under strict liability too, in order to minimize their unit costs. But these unit costs, and thus the price, will equal 15 because the unit costs will include expected liability expenses of 3.//

1.2 Levels of activity. Assume, as is natural, that an increase in a firm's level of

---

[47]As in chapter 2, it is assumed here that the negligence rule works perfectly. If it does not and firms are sometimes found negligent, the price will include a component attributable to expected accident losses (but this component would be less than expected accident losses).

production will result in a proportional increase in expected accident losses, given the firm's level of care. The determination of the socially optimal level of production of firms will then be virtually the same as the determination of the socially optimal level of injurers' activity discussed in the previous chapter. This is illustrated by elaborating the previous example.

Example 2. Because the social costs of production per unit, including expected accident losses, are 15 in Example 1, social welfare will be enhanced by production when and only when a customer obtains utility exceeding 15 from a unit. Suppose, for instance, that there are 5 customers who would derive the utilities shown in Table 2 from purchasing the product. (Or suppose that a single customer obtains increments to utility as shown in the table from purchasing successive units of the product.)

Table 2: Utility from the Product	
Customer	Utility from the product
A	40
B	20
C	17
D	13
E	11

Here, only customers A, B, and C, who derive utility greater than 15, should purchase the product; the optimal level of production (and of consumptions) is thus 3.//

The general point of this example is that it is socially optimal for production to proceed when, but only when, the utility customers derive from consuming additional units exceeds the sum of the direct production costs, the costs of care, and the expected accident losses associated with the additional units.

With this in mind, the analogues to the conclusions from the previous chapter about levels of activity can easily be seen to hold. Specifically, under the negligence rule the level of production will be higher than optimal (and thus too many accidents will occur): Because the price will not include expected accident losses, customers will make purchases when the utility they derive from the product is less than the social cost of production per unit. In Example 2, the price under the negligence rule will be 12. Therefore, not only A, B, and C, but also, undesirably, D will buy the product. (Because, under the negligence rule, the price of gasoline will not incorporate expected accident losses due to its transport, too much gasoline will be purchased.)

Under strict liability, however, the level of production will be optimal, because the price will equal the social costs of production per unit. In Example 2, the price will be 15, so only customers A, B, and C will purchase the product. (Because, under strict liability, the price of gasoline will include expected accident losses, the right amount of gasoline will be purchased.)

The conclusions about victims' levels of activity are also as before. Under the negligence rule, victims will choose their levels of activity optimally, for they will bear their accident losses. And under strict liability with the defense of contributory negligence, victims will engage in their activity to too great an extent, for they will be compensated for their accident losses as long as they take due care. Thus, again, the choice between strict liability and negligence rules will depend on whether it is more important to control injurers' levels of activity -- here firms' levels of production -- than victims' levels of activity.

       1.3 Exclusion of the level of production from the determination of negligence. It was implicitly assumed above that the level of production is not taken into account in the determination of negligence. This assumption describes actual practice -- firms, of course, are never found liable for having produced too much -- and is justified by the fact that if courts were to decide on permitted levels of production, they would have to determine and balance costs of production against consumer valuations. The courts' problem, in other words, would be tantamount to that of devising production responsibilities in a centrally planned economy.

       1.4 Actual liability of firms to strangers. The liability of firms for harm done to strangers is determined as described generally in the last chapter. A distinction must be made, however, between accidents in which harm comes about in the course of productive activity (such as when gasoline explodes during transport) and accidents in which harm to strangers is caused by products after their sale (such as when a boiler is purchased and explodes harming strangers). In accidents of the latter type, the finding of liability is complicated by certain doctrinal considerations and especially by the possibility that the purchasers of products may have played a contributory role.[48]

       Note on the literature. Legal scholars and economists have virtually always mentioned that when strangers might be harmed by firms, imposing strict liability would raise prices and reduce purchases relative to not imposing liability. But a comparison with the situation under the negligence rule, in which prices do not reflect accident losses and are socially too low, has not usually been made. This comparison is first developed in Polinsky (1980c) and Shavell (1980c).

2. Victims Are Customers of Firms

As was indicated at the outset, firms' behavior in this case will be influenced not only by their potential liability, but also by customers' perceptions of product risks, for the latter will affect customers' willingness to make purchases.[49] More precisely, a customer will buy a product only

---

[48]See Keeton, Dobbs, et al. 1984, sections 93, 95-98 on the United States; and Stone 1983, sections 260, 289-91, and Von Bar 1998, 2: 297-311, 418-24, on European countries.

[49]The analysis of liability when victims are the employees of firms would parallel the analysis of this part. In much the same way that customers, if informed, can decide not to purchase unsafe products or can insist on lower prices as compensation for bearing extra risk, employees, if informed, can decide not to work at firms with unsafe working conditions or can demand higher wages as compensation for bearing added risk. I do not examine the issue of firms' liability to employees in part because this would be similar to the present analysis and in part because, in fact, employers tend not to be liable to employees. Employees are generally barred from suing employers by workers' compensation legislation. (This legislation provides that employees may obtain compensation for accidents arising at the workplace and that employers must pay insurance premiums to support the compensation program.) See, for example, Dobbs 2000, 1097-1108; Keeton et al. 1983, chapter 19; and Larson 1994.

if the utility of the product to him exceeds its perceived *full* price -- the price actually charged in the market plus the perceived expected accident losses that liability payments would not cover and thus that he would have to bear. The expected accident losses that a customer perceives that he would have to sustain will depend on his information about product risks. Alternative assumptions about customers' information are considered in the following two sections.

    2.1 Customers' knowledge of risk is perfect. When customers' knowledge is perfect, firms will be led to take optimal care even in the absence of liability. To see exactly why, observe that in the absence of liability customers will bear their losses, and the full price will equal the market price plus expected accident losses. (The full price of a water heater will be seen as its price in the market plus the expected losses due to the possibility that it will rupture.) If a firm were to take less than optimal care, its potential customers would recognize this and factor into the full price the relatively high expected accident losses. Consequently, the firm's customers would go elsewhere; they would prefer to make their purchases from competitor firms exercising optimal care and offering the product at a lower full price, although at a higher market price. In other words, the force of competition will lead firms to take optimal care despite the absence of liability.

    Example 3. Suppose the situation is as in Example 1, except that the victims are customers; and assume that firms do not face liability for accident losses. A firm that does not take care may be able to set the market price of its product at the direct production cost of 10, but the full price will then be 19, for the firm's customers will add to the market price the expected accident losses of 9 that they will bear. The firm would thus lose its customers to firms that do take care. The price charged by firms that take care will be 12 (because the price will have to include the cost of care of 2), yet the full price will be just 15, because expected accident losses will amount to only 3. Hence, a firm that does not take care will not survive in competition against firms that do take care.//

    Firms will also be led to take optimal care under strict liability with the defense of contributory negligence and under the negligence rules. Similarly, customers will be led to take optimal care in their use of products under these liability rules,[50] as well as in the absence of liability.

    Moreover, customers will buy the socially optimal amount of the product regardless of the absence or presence of the foregoing liability rules. This is true because the full price that customers will compare with their utility will not be affected by the absence or presence of such liability. In particular, the market price both in the absence of liability and under the negligence rule will equal the cost of optimal care plus direct production costs, as firms will be led to take optimal care in either case, and customers will add to this market price the expected accident losses in calculating the full price. Under strict liability with the defense of contributory negligence, the market price will simply equal this same full price. In Example 3, the market

---

    [50]Under strict liability without the defense of contributory negligence, however, firms might take excessive care and customers might take inadequate care. For example, users of water heaters might not drain them or watch for signs of leakage even though this would cost users little in time and effort. Manufacturers might therefore be led to produce heaters with safety features that are expensive relative to the cost of users' care.

price will be 12 in the absence of liability and under the negligence rule, and the full price will be 15; under strict liability the market price and the full price will be 15. Thus in all cases only those customers for whom the utility of the product exceeds 15 will buy it, which is the socially desirable outcome.

2.2 Customers' knowledge of risk is imperfect. Suppose now that customers do not have enough information to determine product risks at the level of individual firms. (Customers cannot ascertain the risk of rupture of a particular firm's water heaters.) Then firms will not take care in the absence of liability. No firm will wish to incur added expenses to make its product safer if customers will not recognize this to be true and reward the firm with their willingness to pay a higher price. Liability will thus be needed to induce firms to take optimal care.

Furthermore, the level of care taken by customers will not be optimal in the absence of liability. Customers will take too little care if they underestimate risks and too much care if they overestimate them. In the presence of liability, however, customers who possess accurate knowledge of the level of due care used to determine contributory negligence may be led to take due care despite their misperception of risk.

In addition, the quantity of customers' purchases will not be optimal in the absence of liability or under the negligence rule. If customers overestimate risks, they will overestimate the full price and might decide not to buy products when in fact the utility of the products exceeds the true full price. If customers underestimate risks, the opposite problem might occur; they might make purchases that are not in their interest. Under strict liability with the defense of contributory negligence, however, customers will make appropriate decisions whether to buy products regardless of their misperception of risk. It will not matter that customers incorrectly estimate risks since they will be fully compensated for their losses (because they will take due care); the market price will then reflect the true risk of accident losses, and it will be this market price alone on which customers base their decisions to make purchases.[51]

2.3 Actual customer knowledge of risk. Before I comment on the analysis of the last two sections, it will be helpful to consider briefly the likely character of customers' knowledge of risk.

One point to emphasize is that customers' knowledge of risk will vary with the type of product or service. Customers' knowledge of the risks attending use of a wide class of modern-day products (automobiles, drugs, machines) is, one assumes, limited in significant ways because of customers' quite natural inability to understand how the products function. And customers' knowledge of the quality of most professional services (medical, legal, architectural) is, one supposes, similarly limited. By contrast, customers' information about the risks of common items of fairly simple design (hammers, bicycles, can openers) is probably good on the whole, and the same is likely true of their knowledge of the risks of many of the services that they purchase in ordinary life (barbering, sports instruction).

---

[51]This conclusion would usually be different, however, if losses include nonpecuniary elements for which liability awards would not or could not fully compensate. Suppose, for example, that product-related accidents could result in the death of a customer (as when an automobile's brakes fail). Then a customer would base his purchases in part on his estimate of the risk of death in an accident; if he incorrectly estimated this risk, his decision whether to purchase the product might be inappropriate.

Not only will customers' knowledge of risk vary with the type of product or service, it will vary also with the type of customer. Commercial customers will often have relatively accurate knowledge of risk because they tend to be repeat purchasers, buy in large quantity, and make decisions in a calculated way. The typical consumer may be in a quite different position; he buys many products (especially durables) only on an infrequent basis and may not have the ability or the motive to approach his purchase decision in the manner the commercial customer would.

When customers' knowledge of risk is imperfect, there does not seem to be an appealing general assumption to make about the direction of their errors. A customer's assessment of the risk of a particular product or service will tend to be based on his estimate of the average risk for the class of products or services that have the same outward appearance as the one in question. Because actual risks will deviate from average risks about as often from above as from below, the frequency with which customers underassess risks should approximate the frequency with which they overassess risks -- assuming that they correctly perceive average risks. Of course, customers may not accurately perceive the average, but systematic mistakes in their assessment of risk for a class of products or services can be either positive or negative. Customers can readily be imagined to exaggerate certain kinds of risks, because, for instance, of their vivid aspect (dying in an airplane crash), and they can well be thought to underestimate other kinds of risk, because, say, of the innocuous appearance of the products creating the risks (could drinking hot liquids from styrofoam cups release a carcinogen resulting in stomach cancer?)[52]

That customers' information about risks may sometimes be imperfect seems inevitable. As was suggested above, customers' ability to ascertain risks directly is, , naturally limited by their incomplete knowledge of how products work and by their lay understanding of professional and many other services. Also, their ability to evaluate risks that are numerically small -- as the risks of accidents often are -- may be questioned. The problem that customers may be unable to learn directly about risks may be remedied if customers are apprised of risks by firms. But firms clearly lack appropriate incentives to provide information about the dangerousness of their products and services.[53] In addition, organizations specializing in the collection of information about risks may not be able to earn enough (through sale of publications like *Consumer Reports*) to finance their activities at a socially desirable scale, in part because individual buyers can pass on the information to others in various ways. Finally, the very capacity of customers to absorb and act upon information about the risks they face seems restricted. Customers purchase a great variety of products and services, and the risks of even a single one may be complicated to describe because they depend on the manner and the circumstances in which the product is used

---

[52]Eisner and Strotz 1961 discuss the overestimation of the risk due to airplane crashes. For a general discussion of pyschological factors affecting the assessment of risk, see Kahneman, Slovic, and Tversky 1982 and Tversky and Kahneman 1974; see also Jolls, Sunstein, Thaler 1998 for a synthesis and review of literature. For empirical study of risk assessment, see, for example, Viscusi 1983, 1992, 1998.

[53]This is not to deny that firms may have a legal duty to provide information about risk (liability for failure to warn of defects) or that they can secure marketing advantages or enhance their reputation by doing so. But such motives to provide information about risk arguably are imperfect.

or the service is performed. Customers could not realistically be expected to keep track of and to employ all this information even if it were freely available to them.

With these observations in mind, the reader should assume below that there will sometimes be a useful role for liability to play in reducing risks and in influencing the volume of purchases.[54]

2.4 Problems in applying the negligence rule. Information about firms' conduct and about their products and services may be particularly difficult for courts to obtain or evaluate as they arrive at a determination of negligence.[55] Such information may be of a complicated, technical nature (dealing, for example, with industrial engineering, or with the practice of medicine), or it may be special in character (concerning idiosyncratic features of the production process, or a particular patient's condition), or it may have to do with events that occurred relatively far in the past (the production of an old machine, or the treatment of a patient many years ago). Consider, by contrast, that information will be relatively easy to obtain or evaluate in the typical accident between strangers, as when a person fails to clear a sidewalk of ice and someone slips on it and breaks his leg.

Courts' difficulty in obtaining and evaluating information about firms' conduct leads to two problems. First, courts may be likely to make errors in determining optimal levels of due care. When firms are able to predict courts' incorrectly calculated levels of due care, firms will often be led to take these levels of due care and thus to take excessive or insufficient levels of care, as the case may be. And when firms are unable to predict levels of due care, or when there are other uncertainties surrounding the determination of negligence, firms may well be led to take excessive levels of care so as to avoid being found liable by mistake (a manufacturer may use an undesirably costly safety feature, or a physician may practice "defensive medicine").[56] The second problem is that courts may fail altogether to consider certain dimensions of firms' behavior in the negligence determination, either for want of any evidence or because evidence is scant. With respect to such dimensions of behavior, firms may do little or nothing to reduce risk.

These problems are avoided under the strict liability approach because firms will be motivated to take all justified steps to reduce risk, and only those steps, whether or not courts would be able to decide what steps could and should have been taken.

2.5 Problems in applying the negligence rule to research and development and product design decisions. An important illustration of the problems with the negligence rule concerns research and development with regard to product safety and design. To make a determination of negligence in this area, courts are faced with a complex task: They must decide whether, at the time that a firm had an opportunity to engage in an investigation, the then-relevant probability and value of success were sufficiently high to warrant the costs of the

---

[54]This general view is a theme of Croley and Hanson 1993.

[55]Much of what is said in this and the following sections applies equally in the context of accidents involving firms and strangers.

[56]Why firms may be likely to take excessive care, even though mistakes in the negligence determination may favor them as well as disfavor them, is explained in section 1 of chapter 4.

investigation. Because courts will be prone to make mistakes in determining the probability or value of success or the costs of investigation, firms may be led to make socially undesirable decisions. For instance, a firm that is highly uncertain whether a given degree of research or design effort will later be seen by courts as adequate may decide to engage in research to a socially excessive extent. Or a firm that believes that courts would never learn that it had a particular opportunity to reduce a risk (for example, that a pharmaceutical company had a chance to develop a substitute drug without an adverse side effect) may decide not to pursue the opportunity. Indeed, the likelihood of this outcome is increased by the perversity that the initial pursuit of an avenue of investigation could provide the very evidence that would allow courts to conclude that a research opportunity had existed.

 2.6 Problems in applying the defense of contributory negligence under strict liability. Courts may experience difficulty in determining due care levels of customers and their actual care levels, as well as in incorporating various dimensions of their behavior into due care. It may be hard to determine what customers can and should do to reduce risk (whether users of lawn mowers should wear safety glasses in view of the danger that stones would be thrown up by the cutting blades), how customers actually use products (whether, when mowing, a person tries to steer clear of areas with stones), whether customers adequately maintain products, and the like. In addition, courts typically do not include the intensity of use of products (how often a lawn mower, a can opener, or a forklift is used) in determining customer negligence, because calculation of the appropriate intensity of use is a practical impossibility or because evidence on the actual intensity of use is difficult to obtain.

 As with courts' difficulties in determining the negligence of firms, the courts' difficulties in determining the negligence of customers can lead to two types of problems. With respect to dimensions of behavior that are included in the determination of contributory negligence, customers may be led to do too little or too much to reduce risk, and with respect to dimensions of behavior that are left out of the determination, customers will do too little to lower risk. In any event, the problems would be lessened if, as under the negligence and no-liability regimes, customers expect to bear their losses.

 2.7 Strict liability versus negligence reconsidered. The discussion of the last several sections should help to organize thinking about important factors bearing on the appeal of strict liability versus the negligence rules as means of providing incentives toward safety. To illustrate, suppose that individuals' knowledge of the health risks associated with use of microwave ovens is imperfect (there is a potential need for liability); that the harmful effects of microwave radiation might be of substantial importance (a liability-induced reduction in radiation might prevent significant injury); that the possibilities for changing the design of microwave ovens would be hard for courts to ascertain (a determination of negligence about oven design would be problematic); and that there is relatively little that users of microwave ovens can do to reduce risk (there is no real issue of contributory negligence).[57] In such a case, employment of the strict liability approach rather than the negligence rule would be desirable on grounds of creation of incentives toward safety in product design.

---

[57]The doors of most microwave ovens must be closed for them to operate, so that users cannot "cook" themselves, but arguably users could take care not to stand too close to ovens when they are in operation.

The situation might be different, however, with respect to use of commercial freezers and the risk that they would break down, causing frozen foods to thaw and spoil. Suppose that buyers of the freezers, being in business, know fairly well the risks that freezers would fail (the potential need for liability is small in the first place); that the scope for manufacturers of freezers to reduce risk of freezer failure is modest because the risk is already low (liability could not reduce risk substantially); that the adequacy of user maintenance of freezers -- checking coolant levels, keeping filters clean -- would be difficult for a court to determine in deciding upon contributory negligence; and that the intensity of use of freezers -- the amount and value of frozen food kept in each freezer -- would be omitted from a negligence determination (implying that users would be likely to overuse the freezers unless they bear the losses of freezer failure). Here the negligence rule would be the better form of liability on grounds of creation of incentives.

Of course, in these two examples it was assumed that all the relevant factors worked in the same direction, in favor of either strict liability or the negligence rule; in reality, this will rarely be the case.

2.8 Product warranties. An addition to the model so far considered would allow firms to offer product warranties, which is to say, to choose their own liability rules.[58] The type of warranty that a firm would offer is the one that would minimize the full price of its product as perceived by customers. A firm not offering that warranty would lose its customers to competitors. This means that if customers do not misperceive risks, the warranty that is sold will be the one that results in the lowest true full price and therefore that is socially best. For example, if buyers of commercial freezers have good knowledge of the risks of breakdown, the character of the warranty on the freezers would reflect the optimal balancing of manufacturer and buyer incentives. A warranty that covered the freezer motor but not the coolant system might serve as an implicit inducement for manufacturers to improve the reliability of the motor and for buyers to maintain the coolant system properly. Such an arrangement would result in a lower full price of freezers than if the warranty covered the coolant system as well as the motor.

If customers misperceive risks, however, the warranty that is sold in the market may be socially undesirable. Notably, customers who mistakenly think that a risk is lower than it actually is will tend to buy warranties with terms limiting or disclaiming coverage of that risk. (If customers overestimate risk, they will buy warranties that are too inclusive.) To illustrate, consider the case where customers erroneously believe that a risk is nonexistent. Imagine, for instance, that they believe there is no chance that electric pencil sharpeners will throw off slivers of wood and cause injury when in fact there is this chance. Customers will therefore place no value on a warranty term giving coverage against injury due to such events, although offering the term would cost manufacturers a positive amount. Electric pencil sharpeners will therefore be sold without a warranty term covering injuries due to slivers, or with disclaimers of liability for these injuries. Consequently, manufacturers of electric pencil sharpeners will not have an incentive to reduce optimally the risk of sliver accidents. Moreover, because customers are unaware of this risk, they will buy too many electric pencil sharpeners (as opposed to safer hand-

---

[58]On the economic theory of warranties, see Cooper and Ross 1985, Grossman 1981, Priest 1981, and Spence 1977; see also the survey of Wehrt 2000.

operated pencil sharpeners, or mechanical pencils).

It follows that where customers misperceive risks, it could be socially beneficial for courts to override certain terms of warranties, especially by broadening firms' responsibilities for injury. For courts to know when to override or to expand coverage terms in warranties, however, requires that they be able to distinguish between situations in which customers misperceive risk and situations in which they do not. Courts would have to be able to determine whether customers misperceive the risk that electric pencil sharpeners will cause injury from slivers, for if customers do understand this risk, they might still desire limitation of the warranty for such injuries (perhaps because they can reduce the risk of accident by discarding old pencils that are likely to break apart, perhaps because they have health insurance that would cover losses due to such injuries).

2.9 Imperfect competition and market power. Finally, it should be noted how, if at all, the possibility that firms in a less than perfectly competitive market have the ability to set price above cost affects the conclusions of this chapter.[59] Firms that enjoy such market power will wish to minimize unit costs, as do firms in a perfectly competitive market, in order to maximize profits. Because the conclusions reached above about firms' exercise of care rested only on the assumption that firms seek to minimize unit costs, the statements made about liability and firms' levels of care will not be altered where firms possess market power.

The enjoyment of market power by firms will, however, make some difference with respect to the social desirability of the effect of liability on levels of production. Because firms with market power will set price higher than unit costs, customers will purchase less than they do in a perfectly competitive market setting. Therefore, under strict liability firms with market power will set price above the sum of their production cost, the cost of care, and accident costs per unit; and customers will tend to purchase too little, rather than the optimal amount, of the product due to the higher price. Under the negligence rule, because firms will set price above the sum of production cost and the cost of care per unit, customers' tendency to purchase too much if they underestimate risks will be counteracted, and so forth.

2.10 Actual liability of firms to customers. In most jurisdictions in the United States today, firms are held strictly liable for accident losses caused by defects in their products, though an aspect of negligence is involved in the definition of product defect.[60] Customers need not prove negligence in the production process; they need only show that their losses were due to defects in products. Firms may sometimes avoid or reduce their liability when accident losses were the result of product misuse or other contributory behavior of customers[61] or of dealers, and firms may challenge whether losses were in fact caused by their products. But there has been

---

[59]On market power and the effect of liability rules, see Epple and Raviv 1978 and Polinsky and Rogerson 1983.

[60]The description here is confined to liability of producing firms to customers; it does not include liability of dealers to customers. On product liability in the United States, see for example Dobbs 2000, chapter 24, Phillips 1998, Keeton, Dobbs, et al. 1984, chapter 17, and Shapo 1994.

[61]Although the defense of contributory negligence is often not permitted, use of comparative negligence and doctrines relating to product misuse and to assumption of risk may allow courts opportunity to take into account plaintiffs' behavior.

some narrowing of these defenses and thus an expansion of firms' liability; moreover, firms are increasingly prevented from escaping liability by having disclaimed it in warranties.[62]

In two important areas -- product design and warning of risk -- the negligence rule is employed. Specifically, a firm will be held liable for harms resulting from a dangerous characteristic in all units of its product if an alternative, safer design could have been used at reasonable cost. And a firm will be held liable for failure to warn of a product risk if the firm could have done so at a reasonable cost.

In England, France, and Germany, the trend in product liability has been in the direction of strict responsibility for defects, with this result often being reached by other legal doctrines (presumption of producer negligence where losses are caused by defects, contractual liability, or implied warranties).Yet the scope of product liability does not appear to be as great in these countries as in the United States.[63]

Note on the literature. Oi (1973) and Hamada (1976) examine models of product liability in which victims are customers and possess perfect information about risk; Goldberg (1974) emphasizes that legal intervention must be premised on impefect customer information. Spence (1977) analyzes strict liability in a unilateral model of accidents in which victims are customers who misperceive risk and firms offer warranties. In Shavell (1980c), I consider strict liability and negligence rules in a bilateral model of accidents along the lines presented here. For empirical work on product liability, see for example Higgins (1978), Priest (1988), and Viscusi (1991).[64]

---

[62]See, for example, Epstein 1980, Priest 1991, and G. Schwartz 1992 on the expansion of firms' liability.

[63]See Fleming 1998, chapter 23, on product liability in England; Markesinis 1994, 79-95, on Germany; Zweigert and Kötz 1998, 676-78 on France and Germany; and Stone 1983, sections 257- 95, and Von Bar 1998, 2: 418-24, on Europe.

[64]See also Litan 1991 and the surveys of Dewees et al. 1996, chapter 4, and Geistfeld 2000.

Chapter 4

EXTENSIONS OF THE BASIC THEORY

In this chapter I consider a number of extensions of the basic theory of liability and deterrence. In the first four sections, I discuss various issues concerning the negligence determination, the judgment-proof problem, and vicarious liability; in the later sections, I examine a number of topics about damages, that is, about the magnitude of liability.[65]

1. Problems in the Negligence Determination
Factors leading to uncertainty in the finding of negligence, and the consequences of such uncertainty, will be considered in the initial subsections here. Then the effect of systematic, anticipated error in the courts' determination of due care levels will be analyzed. At the end, the effect of misperception of due care levels will be discussed.

1.1 Uncertainty in the finding of negligence. One factor leading to uncertainty in the finding of negligence is that courts may err in assessing a party's true level of care. For example, a court might not accept a physician's claim that he had performed a diagnostic test (that he listened carefully to a person's heartbeat after a series of exercises) when in fact he had done so. The possibility that a court would make an error of this type might lead a physician to administer a redundant but easily verifiable test (such as an electrocardiogram) that would reduce the chance of a court finding him negligent by mistake. Of course, the possibility that a court would make an opposite type of error may also exist. A court might decide that a physician had taken proper care when in truth he had not. For instance, a court might conclude from incomplete medical records that there was no need for a physician to refer his patient to a specialist when the patient should have seen one.

The significance of these two types of error, however, is not likely to be the same. The disadvantage to a party of being found negligent by mistake is that he will have to pay the victim's losses. This disadvantage will often be of greater importance than the savings that the party could obtain by reducing his level of care somewhat and hoping that he would erroneously escape liability if an accident occurred.

The reader should not be surprised, then, to learn that a fairly general consequence of uncertainty in the assessment of true levels of care is that parties will tend to take more than due care -- and thus to take socially excessive levels of care (presuming that due care is set at socially optimal levels).[66] To illustrate, consider the following example.

Example 1. The probability of an accident that would cause a loss of 100 is related to the

---

[65]The material here draws on chapters 4, 6, and 7 of Shavell 1987a, which also contains proofs of conclusions.

[66]I am not saying that uncertainty in the assessment of care will always lead to excessive care. Obviously, other things being equal, if there is a high enough chance of overassessment of care and a low enough chance of underassessment, parties will take less than due care. But in a wide class of situations (including ones in which the chance of overassessment of care exceeds by a significant amount the chance of underassessment), parties will take more than due care.

level of care as shown in Table 1. The socially optimal level of care, which is assumed to be due care, is moderate care. If there were no chance of mistake in courts' assessment of care, parties could avoid liability for sure by taking moderate care at a cost of 3; they would not take high care, since that would involve a greater cost of 5.

Table 1: Care Level and Accidents				
Level of care	Cost of care	Accident probability	Expected losses	Total social costs
None	0	15%	15	15
Moderate	3	10%	10	13
High	5	9%	9	14

Suppose, however, that there is a 33 percent chance that courts will misperceive care by one level and a 5 percent chance that courts will misperceive care by two levels. That is, there is a 33 percent chance that no care would be seen by courts as moderate care and a 5 percent chance that no care would be seen as high care. Further, there is a 33 percent chance that moderate care would be seen by courts as none and a 33 percent chance that moderate care would be seen as high care. And there is a 33 percent chance that high care would be seen by courts as moderate care and a 5 percent chance that high care would be seen as none.

In this situation, parties will take high care. If they take no care, their expected expenses will be $62\% \times 15\% \times 100 = 9.3$ (as they will mistakenly escape liability $33\% + 5\% = 38\%$ of the time). If they take moderate care, their expected expenses will be $3 + 33\% \times 10\% \times 100 = 6.3$ (for they will mistakenly be found liable 33 percent of the time). Yet if they take high care, their expected expenses will be only $5 + 5\% \times 9\% \times 100 = 5.45$ (because they will mistakenly be found liable only 5 percent of the time).//

As this example shows, if raising the level of care reduces the chance of being found negligent by mistake, parties may decide to take more than due care, even though the chances of courts' overestimating care are as large as the chances of their underestimating care.[67] The example illustrates also the point that despite parties' increasing their level of care, they may still face a positive risk (5 percent in the example) of being found negligent if they cause accidents.

---

[67]Actually, in the example, as long as the chance of overestimating care by one level is less than 58.66 percent -- a chance substantially exceeding the 33 percent chance of underestimating care -- parties will still take high care. Assuming the chance of overestimating care by one level is 58.66 percent, parties who take no care will escape liability 58.66% + 5% = 63.66% of the time, so their expected expenses will be $36.34\% \times 15\% \times 100 = 5.45$; parties who take moderate or high care will expect to spend, as before, 6.33 and 5.45, respectively. Thus taking no care and taking high care will result in equally low expected expenses. If the chance of overestimating care is lower than 58.66 percent, taking no care will result in higher expected expenses.

Much the same conclusions hold with respect to two other factors leading to uncertainty in the finding of negligence. One of these factors is that a party may be unable to control completely his *momentary level of care*. A driver may be unable to control completely his level of care at each instant (because of a lapse of attention, a sudden glare, a sneeze), or a physician may be unable to act with all the care he intends with each of his patients on each of their visits. However, as it is the driver's care at the time of an accident and the physician's treatment of the particular patient on a particular visit that courts will ordinarily consider in determining negligence, the driver and the physician will generally bear some uncertainty regarding their being found negligent. A little reflection should convince the reader that such uncertainty will usually lead parties to try to take more than due care in order to reduce the likelihood that their momentary level of care will fall short of due care (in terms of their usual habits and attitudes -- see the next section) and thus cause them to be found negligent. (The logic behind this assertion is essentially that of the previous paragraphs, that the disadvantage of being found negligent will outweigh the advantage of conserving on the cost of taking care.)

The other factor leading to uncertainty in the determination of negligence is the level of due care that will be applied by courts. It may be difficult for a party to predict how courts will assess the cost of care or its effectiveness in reducing risk, and thus what they will determine due care to be. There may be uncertainty, for instance, in how courts will evaluate the cost to a physician in time and effort of performing a diagnostic test or in how courts will assess the value of the test in providing information about a disease; in this case the physician will not know whether courts will see failure to perform the test as negligence. It should be clear to the reader that such uncertainty will tend to induce parties to take higher than desirable levels of care to guard against being found liable by mistake.[68]

1.2 Remarks on uncertainty. The relative importance of the three sources of uncertainty – courts' errors in assessing true levels of care, parties' inability to control their momentary level of care, and courts' errors in calculating levels of due care -- will depend on the context.[69] For example, when there are few witnesses to, or little evidence concerning, a party's act, errors in assessing true levels of care may be important; when courts are not able to obtain or to evaluate reliably information about the costs and benefits of care, errors in the calculation of the level of due care may be important (a problem that may be of general significance for physicians and other professionals, or for firms using new technology).

With respect to parties' inability to control their momentary levels of care, three comments seem worth making. First, an individual's momentary level of care can be regarded as an imperfect indicator of his true, and inherently unobservable, level of care, namely, the degree

---

[68]The points made in this subsection were first studied by Diamond 1974a and further developed in Calfee and Craswell 1984, Craswell and Calfee 1986, and chapter 4 of Shavell 1987a; qualifications to these points are pointed out in Grady 1983 and Kahan 1989, as observed in note 36 below. The notion of defensive medicine comports with the explanation for excessive care advanced here; for empirical evidence of defensive medicine, see, for example, Danzon 1985 and Kessler and McClellan 1996.

[69]See Tunc 1983, sections 141, 143-46, for a discussion of the importance of uncertainty in the negligence determination in different legal systems.

to which he adopts a *prudential mental attitude*. Hence, in strict logic, the cause of uncertainty in the finding of negligence due to an individual's inability to control his momentary level of care may be viewed as courts' inability to assess an individual's true prudential mental attitude. Second, one wonders whether courts might sometimes lower the level of due care in implicit recognition of parties' problems in controlling their momentary level of care. (Might not courts allow for some irregularity in driving behavior, knowing that individuals cannot maintain full concentration at all times?) Third, there are two types of situations that appear to involve uncertainties similar to those regarding the momentary level of care: situations in which parties are responsible for the negligence of subordinates whose behavior they cannot control completely; and situations in which parties operate machines that occasionally function erratically.

Finally, it should be added that the more general interpretation of the fact that uncertainty in the level of due care may induce parties to take socially excessive care is that uncertainty about the law may lead parties to take socially undesirable steps in order to avoid liability.

1.3 Anticipated errors in the choice of due care. Now suppose that parties know in advance that the level of due care set by courts will be different from the optimal level, and how so. (But, for simplicity, suppose that the courts can correctly measure parties' true levels of care and that parties can control completely their levels of care.)

It might be that parties know that the due care level will be less than the optimal level. This would be true, for instance, when parties know that they will not be found negligent for failure to use a particular safety device despite its low cost and substantial effectiveness in reducing risk. In such a situation parties will obviously not choose to purchase the safety device; they will not take more than due care.

The other possibility is that the parties know that the level of due care will exceed the optimal level (that a safety device will be required despite its high cost and low effectiveness in reducing risk). In this situation parties will take due care unless its level is so high that they are better off acting negligently. In the latter case, parties will take optimal care since they will, in effect, be strictly liable.

1.4 Misperception of the level of due care. Suppose, finally, that parties misperceive the level of due care that courts will apply. Then parties will take the level of care that they believe constitutes due care, unless it exceeds optimal care by so much that they are better off acting in a way they think is negligent, in which case they will take optimal care. Hence, parties who overestimate due care will either take more than due care or take optimal care; those who underestimate due care will take less than due care.

1.5 Comparison with strict liability. The various reasons why uncertainty surrounding the negligence determination may lead to inappropriate levels of care constitute implicit advantages of strict liability. Because under strict liability there is no investigation of the adequacy of an injurer's care level, the problems that would occur in determining negligence are moot, and levels of care will tend to be optimal.

2. Why Negligence Is Found and Implications of Findings of Negligence

2.1 Reasons for findings of negligence. What explains findings of negligence? This

question arises because, as the reader will recall, according to the basic theory of liability presented in chapter 2, parties were never found negligent: It was in an individual's interest to act with due care, and courts therefore always exonerated injurers.

It is evident from the discussion in the last section, however, that there are a variety of reasons why parties may face a risk of being found negligent, including errors in the courts' assessment of care actually taken or in ascertaining the proper standard of due care, inability of individuals to control their momentary behavior, and inability of firms to control the behavior of employees.

Another significant reason for findings of negligence is that parties may not find taking due care worthwhile, and thus will decide definitely to act in a negligent way. We saw that this might be so when parties anticipate an excessive standard of due care. Parties also might not find it in their interest to take due care because they do not have enough assets to pay a judgment or because they think they would escape suit (these possibilities will be discussed in sections 3 and 9 below).

2.2 Significance of findings of negligence. The occurrence of findings of negligence implies that there is an element of strict liability -- of having to pay for harm done -- associated with use of the negligence rule. Hence, many of the conclusions reached above about strict liability carry over to a degree to the negligence setting. For example, the point that under strict liability injurers will take into account the losses their activity creates has relevance under the negligence rule; injurers will take some account of the losses their activity creates because they will face some risk of being found negligent. In addition, the occurrence of findings of negligence will be referred to later, in chapter 5, to explain why injurers should wish to purchase liability insurance against being found negligent.

3. Injurers' Inability to Pay for Losses: The Judgment-Proof Problem

3.1 Dilution of incentives to reduce risk. If injurers do not have assets sufficient to pay fully for the losses they cause -- and may thus be *judgment-proof* -- their incentives to reduce risk by taking care may be inadequate, as they will treat losses that they cause that exceed their assets as imposing liabilities only equal to their assets. For the same reason, injurers' activity levels will tend to be socially excessive and they will contribute too much to risk. These points are more important the lower are injurers' assets in relation to the harm they might cause (in the extreme, if injurers have no assets, they will have no liability-related incentive to reduce risk).[70]

3.2 Significance of dilution of incentives. (1) There are many contexts in which inability to pay for losses plausibly may lead to dulling of incentives to reduce risk. This is so not only for parties with low or moderate assets, but also for parties with substantial assets whose activities pose special risks. (Consider, for instance, even large corporate enterprises and the chance of fires or explosions causing mass injuries, or the possibility that a widely distributed product has toxic or other dangerous properties.) Incentives are particularly likely to be diluted with respect to those actions that would serve primarily to lower the severity or

---

[70]I analyze the nature of the dilution of incentives due to the judgment-proof problem in Shavell 1986; see also Pitchford 1998 and Ringleb and Wiggins 1990.

likelihood of extremely large losses exceeding parties' assets, yet not of small or moderate losses. (Consider the motive of the owner of a nuclear power plant to spend on a safety measure, perhaps an extra concrete shell around the reactor core, that would limit harm only in a catastrophic accident involving rupture of the core and causing losses far greater than the owner's net worth.)

(2) Incentive problems are exacerbated if parties have the opportunity to shield assets, such as when an individual puts his property in a relative's name or when a firm transfers assets to a holding company.

(3) The problem of dilution of incentives is distinct from the problem that scholars and practitioners often identify with injurers' inability to pay fully for losses, namely, victims' inability to obtain complete compensation. This and related issues (concerning, chiefly, insurance) will be addressed in chapter 5.

3.3 Solutions to the problem of dilution of incentives. Several types of social responses to the problem of inadequate incentives to reduce risk are possible, depending on the circumstances.

One possibility is vicarious liability. If there is another party who has some control over the behavior of the party whose assets are limited, then the former party can be held vicariously liable for the losses caused by the latter. This solution will be discussed in section 4 below.

A second possibility is minimum asset requirements. Parties with assets less than some specified amount could be prevented from engaging in an activity. This approach would ensure that parties who do engage in an activity have enough at stake to be led to take adequate care. However, the minimum asset requirement is a blunt tool that could unduly discourage participation in an activity: Suppose that a person who could not meet the asset requirement would obtain a large benefit from the activity and would cause little expected harm, even though his incentives are diluted.

A third response is regulation of liability insurance coverage. This is a somewhat complicated topic, and discussion of it will be deferred until section 7 of chapter 5.

A fourth approach is direct regulation of parties' risk-creating behavior.[71] Thus, for example, a regulatory authority might mandate that milk be pasteurized or that trucks carrying explosives not travel through tunnels. Such regulation could force parties to reduce risks in socially beneficial ways that would not be induced by the threat of liability, due to its dulled effect from the judgment-proof problem.[72] However, a regulatory authority's ability to devise appropriate regulations is limited by its knowledge. I will discuss this issue further in section 5 of chapter 10.

A final 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 a criminal sanction; see chapters 21 and 24.

---

[71]See, for example, Faure 2000, Kornhauser and Revesz 1998, and Menell 1998.

[72]It should be noted that the judgment-proof problem also limits the ability of a regulatory authority to enforce regulations in some circumstances.

4. Vicarious Liability

     4.1 Definition of vicarious liability. The concern here is with the imposition of liability on one party -- the principal -- for some or all of the losses caused by a second party -- the agent. The principal is presumed to have a relationship with the agent that may allow him to observe the agent's level of care and to control it or come to an agreement about it. The reader may wish to think of the principal and the agent as employer and employee, contractor and subcontractor, or parent and child.

     4.2 Vicarious liability increases levels of care and reduces levels of activity if the agent is judgment proof.[73] Suppose that the agent is judgment proof. Then, as discussed in section 3, the agent's incentives to take care will be inadequate if he alone is liable. However, imposition of vicarious liability alters the situation because it puts the principal's assets at stake.

     If the principal can observe and control the agent's level of care, then imposition of vicarious liability will induce the principal to compel the agent to exercise optimal care, because that will reduce the expected liability payments of the principal. (If the principal is in a contractual relationship with the agent, as would be true of an employer and employee, the principal might have to pay the agent to take added care, but the principal will still prefer to do that in order to reduce his liability payments if the level of care is optimal.)

     If the vicariously liable principal cannot observe and directly control the agent's level of care, the principal will attempt to induce the agent to take care by instituting penalties for adverse outcomes, such as demotion, discharge, or suit in the case of employees. However, since the agent's assets are less than the losses he might cause, the principal will not generally be able to induce the agent to choose the optimal level of care.

     Whether or not the principal can observe and control the agent's level of care, imposition of vicarious liability will lead the principal to reduce the agent's participation in, and level of, risky activity, assuming that the principal can control the agent's level of activity.

     4.3 Factors bearing on the appeal of vicarious liability. (1) The advantage of vicarious liability in desirably affecting incentives to reduce risk will be greater the lower the agent's assets are, and the higher the principal's assets are, relative to the probable magnitude of harm the agent can cause. The advantage will also be greater the better able the principal is to control the agent's behavior.

     To illustrate the relevance of these factors, consider the important example of the large firm and its employees.[74] It is apparent that the assets of employees are likely to be much lower than the losses they could cause, for the scope of a firm's activities is frequently such that a

---

[73]Sykes 1984 contains a general analysis of vicarious liability stressing, in addition to some of the issues of incentives discussed here, the allocation of risk between agents and principals; see also Kornhauser 1982. For analysis of a particular form of vicarious liability, imposed on suppliers of services (such as lawyers, accountants, and lenders) to possibly judgment-proof parties, see Kraakman 1986 and Pitchford 1995. For analysis of another important form, imposed on owners of a firm, see Hansmann and Kraakman 1991 and Halpern 1998.

[74]The reader should be reminded that I am considering here only issues of incentives. In particular, I am not considering any advantages (or disadvantages) vicariously liable parties may have as risk bearers.

single employee's behavior may result in harm to many parties or otherwise lead to significant losses. Thus, were employees only individually liable, one suspects that their incentives to take care would often be seriously inadequate. One also supposes that imposition of vicarious liability on the firm helps to cure this problem because the firm's assets are usually much greater than those of any of its employees and because the firm typically has the ability to exert significant control over its employees' behavior.

In other contexts, and especially those in which a principal engages an agent on a one-time basis, vicarious liability is not necessarily as effective, because there is no natural presumption that can be made about the agent's assets relative to the principal's. Contrast the following two examples: (a) A homeowner (the principal) of average means pays a national pest-control firm (the agent) to carry out extermination services; here the principal's assets are much smaller than the agent's. (b) A large construction firm (the principal) subcontracts with a small, family-owned plumbing company (the agent) to help on a job; here the principal's assets are much greater than the agent's. These examples serve also to illustrate that the principal's ability to control effectively the agent's behavior may be adequate in one situation but poor in the next -- whereas the construction firm should be able to watch over the plumbing subcontractor quite well, the likelihood that the homeowner can judge the performance of the exterminator is not great. Evidently, then, the attractiveness of vicarious liability will depend significantly on the features of the situation at hand.

(2) The desirability of vicarious liability is enhanced by two additional factors. First, principals may have better knowledge than agents about the nature of risk or be able themselves to take actions that can reduce risk. Where this is so, imposition of vicarious liability will obviously lead principals not only to have agents take appropriate care, but also to take additional actions of their own to reduce risk. In the presence of vicarious liability, firms will be led to issue instructions, organize the conditions of the workplace, schedule operations, select employees, and so forth, in ways that better reduce risk. Second, principals may have more information than courts do about the appropriateness of agents' behavior. If that is the case, and if under vicarious liability principals frequently replace courts as the discipliners of agents, fewer mistakes will be made and better conduct will be promoted.

(3) There are, however, disadvantages of vicarious liability that should be kept in mind. Specifically, imposition of vicarious liability will increase the administrative cost of using the legal system, because it will raise the number of defendants named in actions brought by victims, otherwise complicate proceedings, and also engender claims by principals against agents.

4.4 Actual use of vicarious liability. Vicarious liability is a significant feature of legal systems today.[75] Most important, firms are held responsible for the losses caused by their employees. Vicarious liability is in addition sometimes imposed on principals for losses caused by their agents (here I am using these terms in their legal sense), on automobile owners for accidents caused by those whom they allow to drive, on parents for harms resulting from their children's acts, on teachers for their students' negligence, and so forth. Although the pattern of

---

[75]See Dobbs 2000, chapter 22, Fleming 1998, chapter 19, and Keeton, Dobbs, et al. 1984, chapter 12, on Anglo-American law; and Le Gall 1983, Eörsi 1983, and Von Bar 1998, 1: 351- 63, for comparative treatments.

use of vicarious liability is complicated and varies among legal systems, the general principles that are applied seem to be such that the greater the degree of one party's control and authority over a second party, and the more knowledge the first party has about the second party's behavior, the more likely the first party is to be held responsible for the losses caused by the second party.

5. Damages and the Level of Losses

5.1 Damages equal to the level of losses. It was shown above, in chapters 2 and 3, that the threat of liability generally leads parties to take optimal levels of care, and sometimes to choose optimal levels of activity. The arguments for these optimality conclusions were made using examples in which there is one possible level of harm that results if an accident occurs and in which the magnitude of liability -- so-called *damages* -- equals that level of harm. These optimality conclusions carry over to situations in which there are multiple possible levels of harm. Consider the following illustration, in which there are two possible levels of harm.

Example 2. If injurers do not take care, the probability of an accident will be 10 percent. And if an accident occurs, there will be a loss of 100 with a probability of 80 percent and a loss of 500 with probability 20 percent. Expected losses if care is not taken will therefore be $10\% \times (80\% \times 100 + 20\% \times 500) = 18$. Hence, if exercising care eliminates the possibility of an accident and costs less than 18, it will be socially desirable to do so.

Suppose that liable injurers must pay for the losses that occur, whether these are 100 or 500. Then under strict liability injurers will bear expected liability of 18 if they do not take care, and thus will be led to take care when they ought to. Injurers will behave in the same way under the negligence rule if the cost of taking care is less than 18 because they will be liable if they fail to take care.//

This example illustrates why, in the general setting in which an accident can result in more than one level of harm, injurers will act optimally if damages equal actual harm. The reason is that if a liable party must pay for the actual harm he causes, whatever the level of harm happens to be, his expected damage payments will equal the expected harm he causes. And the condition that expected payments equal expected harm is exactly the assumption on which the arguments about optimality of parties' behavior under liability rules has been based in previous discussion. If damages tend to fall short of harm, so that expected payments are below expected harm, incentives to reduce risk will be inadequate, and if damages exceed harm, so that expected damages exceed expected harm, incentives to reduce risk will be too high.[76]

5.2 Actual magnitude of damages. The starting principle in most legal systems is that a liable party should pay for the actual level of losses caused, whether they be high or low. It is said, for instance, that an injurer takes his victim as he finds him, that the injurer should pay for harm caused if the victim turns out to be affected by a latent aggravating condition (a thin skull,

---

[76]It should be noted, however, that if damages equal the expected harm regardless of the actual magnitude of harm, incentives to reduce risk will also be appropriate. This setting of damages, however, would usually not be desirable, for actual harm is generally easier to determine than expected harm (where that is not so, damages may be set equal to expected harm -- see section 7). Moreover, it would mean in particular that damages would be less than actual harm when actual harm is higher than average, something that victims would view as problematic.

hemophilia), and that the injurer should pay similarly if the property he damaged had some structural weakness or was unusually valuable. At the same time, a liable injurer is responsible only for small losses if only a small harm resulted from his act.[77] (Nevertheless, there are subsidiary principles, which will be noted in section 6, under which liable injurers do not pay damages equal to losses.)

6. Damages and the Probability of Losses

6.1 Optimal damages are unaffected by the probability of losses. The conclusion from the last section that behavior will be optimal if the magnitude of liability equals the actual level of losses suggests that behavior will not be optimal if liability is adjusted on the basis of other factors -- in particular, if damages are lowered because a loss was very unlikely, or if damages are raised because a loss was very likely. These two possibilities will now be considered.

6.2 Limitation of damages for unusual losses. Suppose that damages are limited to an average or typical level when losses happen to be unusually high. Then expected liability payments will be less than expected losses and the incentive to take care may be inadequate. In Example 2, if liability for the relatively unlikely losses of 500 is limited to 100, then an injurer's expected liability if he does not take care will be only $10\% \times 100 = 10$, which is less than the expected losses of l8. An injurer will therefore take care only if the cost of doing so is less than 10, rather than whenever the cost is less than 18.

To better appreciate the conclusion that, under present assumptions, liability for unusual accidents should not be limited, observe that the contrary conclusion would lead to a reductio ad absurdum. Any accident, after all, can be seen as extremely unlikely if it is described in sufficient detail. For example, the initially likely-sounding accident in which a person drives his automobile into his neighbor's picket fence becomes a very unlikely one when it is mentioned that the accident occurred on a Tuesday at 4:23PM, that the left side of the automobile's fender struck the fence, and that the eighteenth through twenty-seventh pickets were broken. Were one to contend that liability for unlikely accidents should be limited, one might thus be led to say that liability for any accident whatever should be limited.

Another way to understand this point is to recognize that the magnitude of expected losses reflects all manner of possible accidents (striking the neighbor's fence with the right side or the middle part of the fender if not with the left side, breaking the nineteenth through twenty-eighth pickets if not the eighteenth through twenty-seventh). Expected losses are a probability-weighted aggregation of losses that can arise in many individually unlikely ways. Were liability reduced because of the improbability of the particular accident, expected liability could not equal expected losses.

6.3 A qualification. Nevertheless, it might be acceptable to limit liability for certain accidents: If the possibility of some type of accident is overlooked, then there would be no decrease in injurers' incentives caused by reducing liability for that type of accident. (Note that this argument is not an affirmative reason for reducing the magnitude of liability; it says only

---

[77]See for example Dobbs 2000, 464-65 and Fleming 1998, 234-37, on Anglo-American law, and Stoll 1983, sections 26-28, and Von Bar 1998, 2: 156-62, on law in other countries.

that reducing liability may not have a detrimental effect on incentives.)

To decide which types of accident may be overlooked, consider that individuals cannot practically contemplate each and every one of the multitude of possible accidents that could follow from their actions. People must amalgamate potential accidents into a relatively small number of categories, assign probabilities to the categories, and make decisions with reference to them. In the process, the possibility arises that some accidents will not be taken into account because they do not fit into the list of categories used in decisionmaking. Such accidents might be described as freak.[78]

There are, however, several problematic aspects of a policy limiting liability for accidents whose possibility is overlooked. First, this policy invites parties to deceive courts about accidents that they had in fact contemplated; second, the policy reduces parties' incentives to consider the full range of consequences that could result from their actions; and third, it may increase the costs of adjudication.[79]

6.4 Actual damages for unlikely losses. Although, as was stated in section 5.2, the starting principle in most legal systems is that liability equals losses caused, there are exceptions to the principle in Anglo-American law that fall under the rubric of unforeseeability.[80] According to this notion, liability should not extend to harms that the injurer could not reasonably foresee. It appears from examination of cases and legal commentary that the accidents that are held not to be reasonably foreseeable generally have a far-fetched, extraordinary, freak character.

6.5 Damages for highly likely losses. If liability for accidents that are very likely to occur exceeds the level of actual losses, then parties may have too great an incentive to reduce risk. Suppose that a construction firm that blasts to excavate a large area will be virtually sure to cause some losses. If the firm has to pay for more than these losses because of their high likelihood, the firm may be led to take excessive precautions in blasting, or may decide to use alternative, more expensive means of excavation, even where blasting is socially best.

6.6 Actual liability for highly likely losses. Again, because the starting principle is that liability should equal actual losses, the fact that the probability of losses might have been high does not ordinarily result in liability greater than the losses. A partial exception occurs when the injurer was reckless or knew harm would occur, especially when intended; see section 9.

7. Damages and Courts' Uncertainty about the Level of Losses

7.1 Damages equal to expected losses. Suppose that courts are not able to assess

---

[78]As in the celebrated American case of Palsgraf v. Long Island R.R., in which a package containing fireworks was dislodged from the arms of a man boarding a train and fell under the train's wheels, causing an explosion that knocked a scale onto and injured a woman at a platform some distance away.

[79]Limiting the magnitude of liability will tend to increase the cost per case brought, but it may also reduce the number of cases that are brought and thus lower the costs of adjudication.

[80]On unforeseeability, see Fleming 1998, 237-46, Keeton, Dobbs, et al. 1984, section 43, and the *Restatement (Second) of Torts* 1965, section 435. In other legal systems there are some parallels with unforeseeability (chiefly in the interpretation of the "adequacy" theory), but there does not seem to be an exactly corresponding notion; see the discussion in Honoré 1983, sections 91-93, Markesinis 1994, 107-08, and Zweigert and Kötz 1998, 601-02, 621.

accurately the level of losses that occur in individual accidents, but use estimates of losses that are correct on average. Then liable parties' expected damage payments will still equal the expected losses they cause, so they will still be led to act optimally under liability rules. In Example 2, for instance, suppose that courts are not able to assess whether losses are 100 or 500 but know that expected losses are 180 and impose damages of this amount. Then injurers will be led to behave optimally.[81]

  7.2 Comments on courts' uncertainty. Courts may be uncertain about the level of losses when the harm has already occurred, as when a house burns and the value of its contents are not easy to determine. Similarly, uncertainty may arise when some elements of the harm will occur in the future. A primary example of the latter is when an individual's subsequent earnings will be reduced due to an injury. The interpretation of the point under discussion in this type of situation is that if courts' estimates of future losses are correct on average, injurers will have appropriate motives to reduce risk. There will be no need for courts to determine what victims' losses turn out to be.

  It should also be noted that estimating uncertain elements of harm may be a difficult and much disputed process, raising the administrative costs of use of the legal system. Therefore, it may be socially desirable for courts to exclude uncertain components of losses from the computation of damages if the probable magnitude of these losses, and the consequent dilution of incentives to reduce risk, is not too large. However, an alternative and superior approach is for courts to approximate uncertain components by means of some easily applied formula, and not to allow dispute over this part of the damage calculation.

  7.3 Actual determination of damages in the face of uncertainty over the level of losses. When losses involve harm to property, courts ordinarily attempt to estimate uncertain components that are not too speculative; they follow the same course when property losses are associated with forgone profits, but their approach is conservative.[82]

  With respect to accidents resulting in injuries to persons, courts can usually ascertain medical expenses borne and income lost up to the time of trial fairly reliably, but subsequent medical expenses and diminution in earning capacity may be highly uncertain. In many countries courts estimate these amounts as best they can, often using actuarial and statistical data to award a lump sum to injured parties or, in the case of fatal accidents, to their dependents for loss of support. However, in Germany, the preference is against lump-sum awards; instead, courts favor awards for reduced earnings to be paid on a periodic basis only as long as injured parties actually suffer them.[83]

---

  [81]I am continuing to assume here, as in sections 5 and 6, that when the injurer chooses his level of care and activity level, he does not know what the magnitude of accident losses will be (100 or 500 in the example), he only knows the probability distribution of accident losses. Therefore, the injurer will have proper incentives to take care as long as his expected liability equals expected harm.

  [82]The description here is based mainly on Fleming 1998, chapter 10, sections 1, 3, and on Stoll 1983, sections 12-48, a general comparative treatment of damages and other remedies.

  [83]See McGregor 1983, sections 49-52, and Von Bar 1998, 2: 197-200.

8. Damages and Pecuniary versus Nonpecuniary Losses

      8.1 Definition of the two types of losses. *Pecuniary* losses are those that either are monetary or are losses of goods that can be purchased in markets, in which case the measure of the losses are the replacement costs. *Nonpecuniary* losses correspond to the losses in utility suffered when irreplaceable things have been destroyed, such as family portraits or other unique objects, or, importantly, injuries involving individuals' health, physical integrity, or emotional well-being.[84]

      8.2 Damages equal to the sum of pecuniary and nonpecuniary losses. Because both pecuniary and nonpecuniary losses reduce social welfare, it is clear that parties will be led to act appropriately under liability rules only if damages equal the sum of pecuniary and nonpecuniary losses. If damages do not fully reflect nonpecuniary losses, parties' incentives to reduce risks may be inadequate.

      8.3 Courts' ability to assess pecuniary and nonpecuniary losses. Because pecuniary losses are equal either to actual losses in wealth or to the cost of replacing goods, such losses are often easy for courts to determine. By contrast, because nonpecuniary losses cannot be observed directly, they are difficult for courts to estimate. Hence, it might be thought that courts ought not to attempt to estimate nonpecuniary losses if they are probably small: In this way administrative costs would be avoided, while incentives to reduce risk would be little affected, as was generally suggested in the case of courts' uncertainty about the level of losses. But, as was noted, in principle a better approach to adopt when nonpecuniary losses are likely to be small, and thus not worth the administrative costs of measuring, is to make use of simple tables or formulas. In any case, if nonpecuniary losses are likely to be large, it is important for courts to attempt to estimate them, and especially when pecuniary losses are small. Otherwise, incentives to reduce risk may be seriously compromised. This may be the situation, for example, with respect to the death of young children.[85]

      8.4 Actual liability for nonpecuniary losses. The categories of losses typically described as nonpecuniary include pain and suffering, emotional distress, and the like. Losses of money and of goods for which substitutes can be bought are regarded as pecuniary, with the prices of the substitutes measuring the losses.

      The willingness of courts to increase awards on account of nonpecuniary losses varies considerably among legal systems. French law may be the most liberal in the types of nonpecuniary losses recognized. Anglo-American law is less liberal (though the size of awards,

---

[84]To be more precise about the definitions, consider a model in which there are two goods -- a good that can be directly consumed and from which other goods can be produced, and an irreplaceable good, which cannot be produced. In this model, assume that the utility of an individual equals the number of units of the first good and its equivalent in produced goods, plus the utility to him of the irreplaceable good if he possesses it.

[85]As will be suggested below, the nonpecuniary component of awards for the death of children is low or nonexistent. The pecuniary awards are usually small as well, because they are based on the future loss of support -- often zero -- that parents will suffer in their later years. (For these reasons, one might wonder about the adequacy of the incentives of, say, toy manufacturers to reduce the risk of fatal accidents.) See McGregor 1983, sections 253-55, 273.

when given, seems highest, especially in the United States). German law is more restrictive than French or Anglo-American in the types of nonpecuniary losses for which damages are allowed.[86]

Moreover, treatment of nonpecuniary losses within legal systems may depend on the type of accident and other factors. Nonpecuniary losses are not usually compensated unless there is accompanying physical injury (but if the injury results in death, nonpecuniary losses suffered by the victim's family are not ordinarily compensated, at least in Anglo-American law). Additionally, nonpecuniary losses associated with losses of unique objects of property are rarely compensated.

9. Damages Greater than Losses: Punitive Damages

9.1 Terminological note. It is conventional to refer to damages that are greater than losses as *punitive*, and I shall sometimes use this terminology here. The explanation for the terminology is that damages exceeding losses are often imposed as a form of punishment (see section 9.7), even though, as will be discussed, there are rationales for the imposition of damages exceeding losses that would not naturally be described as punitive in nature.

9.2 Damages greater than losses generally create excessive incentives to reduce risk. As was stated in section 5, if damages are set equal to losses, incentives to reduce risk will generally be desirable. If damages exceed losses, levels of care will tend to be excessive, and levels of activity will be too low (at least under strict liability). Nevertheless, there are several possible rationales for imposition of damages exceeding losses, including the following.[87]

9.3 Escape from suit. Suppose that injurers who ought to be liable might escape suit. This could be so because it is difficult for victims to identify who injured them (as in the case of a driver who flees from the scene of an accident that he caused, or of a firm that discharges an untraceable but disease-causing pollutant into the air). In addition, suit might not be brought because of litigation costs.

If injurers who sometimes escape suit are made to pay only the usual level of damages on those occasions when they are sued, then their expected payments will be less than the expected losses they generate. Consequently, their incentives to reduce risk will be inadequate. For incentives to be restored appropriately, damages must be raised above the level of losses when they are found liable. Specifically, if damages equal losses multiplied by the inverse of the probability of suit, then expected damages will equal expected losses and incentives will be correct. If, for instance, the probability of suit is 50 percent and losses are $10,000, damages should be multiplied by 1/.5 or 2, for then damages will be $20,000, so that expected damages will be 50% × $20,000 = $10,000 the correct amount; and if the probability of suit is 33.33 percent, then damages should be multiplied by 3, making them $30,000 and expected damages

---

[86]See Fleming 1998, 266-70, 285, for a description of Anglo-American law; Stoll 1983, sections 35-48, and Von Bar 1998, 2: 69-88, 169-89, for a general comparative survey; and McGregor 1983, sections 35-47, 146-72, 212-17, 264-73, for a comparative treatment of damages for personal injury and death.

[87]On economic analysis of punitive damages, see generally Cooter 1989 and Polinsky and Shavell 1998b. For empirical work, see, for example, Eisenberg et al. 1997 and Karpoff and Lott 1999.

$10,000; and so forth.[88]

9.4 Illicit utility from causing harm. Suppose here that an injurer will obtain utility from causing losses and that this utility is not credited in social welfare. When, for instance, a person breaks the windows of his neighbor's house because he positively enjoys the unhappiness his neighbor will experience, society may not want to count the person's utility as an addition to social welfare.[89] (Contrast this situation to one in which an individual playing catch accidentally throws a ball that breaks a window; here his pleasure derives from playing catch, and breaking a window does not raise it.)

Given the assumption that the utility the injurer obtains from doing harm is not credited in social welfare, society wants to discourage the injurer's harmful act. To accomplish that, the damages that are imposed must exceed the utility that the injurer would obtain from his act. Therefore, damages may have to be higher than the losses caused. To discourage a person from spitefully breaking his neighbor's windows may require a penalty greater than the cost of replacing the windows because the person may derive substantial pleasure from committing this act.[90]

9.5 Encouraging market transactions. In some circumstances it is possible for a party to communicate with a potential victim before causing harm. For example, this might be so when a firm contemplates infringing on another's copyright. When prior communication is possible, a potential injurer could negotiate in advance with the potential victim to purchase the right to engage in the loss-creating conduct. The firm deliberating about the copyright violation could secure a license to use the copyrighted material.

In such circumstances, it may be socially desirable to induce a potential injurer to bargain and purchase the right to engage in harm-creating conduct; this can be accomplished by threatening to impose punitive damages on the potential injurer if he acts to cause harm.[91] Specifically, suppose that the usual level of damages would be an underestimate of harm for

---

[88]More precisely, if injurers are risk neutral and damages d equal h/p, where h is harm and p is the probability of suit, then expected damages will be $p(h/p) = h$, and incentives will be correct. If injurers are risk averse, then damages need not be as high as h/p to give injurers adequate incentives to reduce risk.

[89]I examine this assumption because it seems to be one that is held by many individuals, and the view that different sources of utility might count differently in the social calculus has a distinguished pedigree -- see for example Mill 1861, 56, and Harsanyi 1977, 62. However, I myself do not find the trumping of malevolent (or other) preferences appealing, for reasons that are articulated in Kaplow and Shavell 2002c, section B.3, chapter 8. In brief, I see no principled basis for distinguishing among sources of utility to individuals, and any way of so doing leads to the possibility that all individuals would be made worse off.

[90]A similar rationale concerns illicit disutility associated with the exercise of care. Suppose that some people claim that the exercise of care would cause them special and peculiar disutility that is not counted as a cost in the social welfare calculus. For instance, a driver might claim that to be attentive to the road is extraordinarily bothersome. If society does not credit this disutility, it will want to induce the person to be attentive to road conditions, and that may require a level of damages exceeding losses.

[91]This point apparently originated with Calabresi and Melamed 1972 and was further developed by Biggar 1995, Haddock, McChesney, and Spiegel 1990, Kaplow and Shavell 1996b, and Landes and Posner 1981b.

some reason (see section 7.2). A potential injurer then might cause harm when doing so is socially undesirable C because the benefit to the injurer might be greater than the low estimate of harm but less than the actual harm. There may be additional undesirable repercussions from an underestimate of harm. If injurers can take property from victims without having to pay its full value, injurers will devote effort to identifying and appropriating such property (copyright violators will seek out material to copy), and victims will expend effort to protect their property (copyright owners will invest resources in preventing duplication of their material). Such efforts are socially wasteful. The foregoing problems can be avoided if punitive damages are imposed for unilaterally causing harm, for that will induce bargaining and exchange only if the injurer's benefit exceeds the property owner's loss.[92]

Another possible reason to employ punitive damages to encourage bargaining and market transactions concerns administrative costs. If compensatory damages are used alone, harm and the taking of property might be more frequently mediated through the legal system by the bringing of lawsuits than if punitive damages are used as well; thus, if bargaining is less expensive than litigation, administrative costs will fall as a result of the use of punitive damages.

9.6 Punishment. A further rationale for punitive damages derives from a consideration that will be addressed later in this book (see section 3 of chapter 23, on the retributive motive), namely, the objective of imposing proper punishment on a wrongdoer. Assume that, given the degree of a party's blameworthiness, individuals believe that there is a correct level of punishment, and that either higher or lower punishment detracts from the utility that individuals obtain from satisfaction of the punishment objective. (This utility is distinct from any utility that individuals obtain from deterrence of harm caused by imposition of penalties.) Acts that have certain outrageous qualities may call for levels of damages higher than losses to help satisfy the punishment objective and thus can justify punitive damages.

Two observations should be made about the punishment objective. First, the optimal level of damages will be an implicit compromise between this objective and deterrence, as discussed above. Second, in considering the punishment objective when the defendant is a firm, one presumes that the goal is to punish culpable individuals within the firm (not the firm as an abstract entity). This means that account needs to be taken of the degree to which blameworthy individuals within firms will be punished as a consequence of imposition of damages on the firm, and also of the degree to which those not considered responsible (employees generally, and perhaps stockholders) will be punished.

9.7 Actual award of punitive damages. Punitive damages are awarded in the United States in cases where parties acted with ill will, malice, or conscious disregard for others, or

---

[92]An important qualification to the argument of this paragraph concerns the question why the use of punitive damages should be expected to induce bargaining, for it might be expected to occur anyway. To illustrate, suppose that punitive damages are not employed and that the copyright holder is willing to pay more to prevent infringement than the potential infringer values infringement. One might expect the copyright holder to pay the potential infringer not to infringe. Thus, punitive damages would not be needed to induce bargaining -- it would occur anyway. However, if there are multiple potential infringers, the copyright holder would be unwilling to bargain with each and every one, so punitive damages would be needed to induce bargaining and prevent infringment by an infringer unwilling to pay enough to satisfy the copyright holder. This point is emphasized in Kaplow and Shavell 1996b.

where their behavior was outrageous or provoked indignation for some other reason, although it is not very common in cases where harm has been done to property alone. In other countries, analogues to punitive damages are sometimes awarded, but are much less important than in the United States.[93]

10. Damages and Victims' Opportunities to Mitigate Losses

10.1 Victim actions to mitigate losses. In the event of an accident, the victim may be able to limit the harm by taking various actions. For instance, if workers from the phone company accidentally cause a telephone pole to fall, and it breaks through the roof of a person's home, the homeowner may be able to keep his losses to a minimum by removing articles from his attic that could be damaged by rain and also by having his roof repaired promptly. (Such actions that victims can take to mitigate losses after accidents occur are, of course, to be distinguished from the precautions that victims can exercise before accidents occur in order to reduce the likelihood or severity of harm.)

10.2 Social welfare optimum. Given the goal of minimizing total social costs associated with accidents, including here the costs of mitigation of losses, it will be socially desirable for a victim to act to mitigate losses if the cost of so doing is less than the reduction in losses thereby accomplished.

Example 3. If a victim takes an action to mitigate losses due to an accident, the losses will equal 100; otherwise losses will equal 150. It will therefore be socially desirable for him to take the action if its cost is less than 50. If, for instance, the costs of the mitigation action is 10, the action should be taken.//

Thus, in the case of the homeowner who can easily mitigate losses by removing articles from his attic, we might interpret 10 to be the cost of removing the articles, and 100 as the loss given that he does so.

If an accident occurs, total social costs due to it should be regarded as the optimally mitigated level of losses plus the costs of mitigation. In Example 3, therefore, the total social costs due to an accident are 110, for if an accident occurs, not only are 100 in direct losses suffered, but also costs of 10 are incurred in mitigating the losses (preventing them from rising to150). Because total social costs due to an accident are 110, it is this figure that should determine the optimal level of care that injurers exercise to prevent accidents from happening.

10.3 Optimal damages. The claim here is that the optimal level of damages is the level of losses that a victim would sustain had he optimally mitigated his losses -- whether or not he actually did so -- plus the costs of optimal mitigation actions. In the example above, therefore, the optimal level of damages is 110. If this is set as the level of damages, the assertion is that two things will be true: Victims will be led to optimally mitigate their losses, and injurers will be induced to choose the optimal level of care in order to prevent accidents.

To explain, consider the victim in Example 3. He knows that he will receive 110 in damages whether or not he mitigates losses, and in particular, he will not receive 150 if he fails

---

[93]See Dobbs 2000, Keeton, Dobbs, et al. 1984, section 2, and Fleming 1998, 2, 23, 27, 562-64, on Anglo-American law, and Stoll 1983, sections 103-25, and Von Bar 1998, 1: 627-31, on law in other countries.

to mitigate losses. Thus, the victim will decide to spend the 10 to mitigate his losses: If he does so, he will sustain true losses of 100 and mitigation costs of 10, so he will be fully compensated, whereas if he does not mitigate, he will lose 40 (that is, 150 in true losses less 110 in damages received). Because, then, victims are induced to mitigate losses optimally, the damages paid by injurers equal the social costs incurred, so that injurers will indeed be induced to take optimal precautions.[94]

10.4 Actual law regarding mitigation. In the Anglo-American and French legal systems the size of awards is restricted to losses that an injured party could not reasonably have avoided plus expenses reasonably incurred in so limiting losses. Thus awards are based on the assumption that any reasonable repairs to damaged property have been made, that injured parties have obtained proper medical treatment to alleviate their condition (but not that they have submitted to dangerous procedures), and so forth. In the German legal system, as well as in some others, the problem of mitigation of losses is viewed as an aspect of the injured party's contributory fault. Since liability is normally reduced on account of contributory fault, the end result is apparently similar to that under Anglo-American and French law.[95]

11. Causation

11.1 Definition and introduction. The principal meaning of causation that will be employed here is the one used in ordinary language: we will say that a person's act caused harm if the harm would not have occurred had the person not committed the act. For example, a person's speeding will be said to have caused an accident if the the accident would not have occurred if he had not been speeding. This notion of causation is sometimes referred to as *causation in fact,* or as *but for causation* (as in "but for the speeding, the accident would not have occurred"), or as *necessary causation,* to distinguish it from other concepts of causation that fall under the heading of *proximate causation.*

A fundamental characteristic of liability law is that a party must have caused harm in order to be held liable for it.[96] The main question to be examined here is how this feature of law affects the functioning of the liability system under strict liability and the negligence rule. Then I will discuss proximate causation and uncertainty over causation.

11.2 Strict liability. As was elaborated in chapter 2, if parties are held strictly liable for harm, they will generally be led to choose the socially desirable level of care and the socially desirable level of activity. The issue of causation, however, was not raised there, and once considered, it becomes evident that for parties to have socially correct incentives to take precautions and to engage in activities, they must escape liability for harms that they do not

---

[94]Wittman 1981 first suggested that both victims and injurers will act desirably if damages equal optimally mitigated losses plus optimal mitigation costs.

[95]See Dobbs 2000, sections 203-05, Fleming 1998, 285-87, and McCormick 1935, chapter 5, on Anglo-American law, and Stoll 1972, section 155, and Von Bar 1998, 2: 562-63, on law in other countries.

[96] See, for example, Dobbs 2000, chapter 9, on Anglo-American law, and Honoré 1983, sections 15, 106-18, and Zweigert and Kötz 1998, 601, 621, on the law in civil law countries.

cause.

To illustrate that in order for the level of care to be optimal, parties must not be liable when they are not the causes of loss, suppose that a firm's production generates pollution, which can cause house paint to peel and will lead to aggregate repainting costs of $100,000. Suppose too that peeling may also come about from prolonged exposure to the sun and that the cost of repainting homes for this reason is $80,000. It is then socially desirable for the firm to invest in a device like a smoke scrubber to eliminate the pollution if and only if the device costs less than $100,000, for $100,000 is the increase in harm due to the pollution. Further, this is precisely how the firm will be motivated to act if it is liable for repainting homes if and only if its pollution is the cause of peeling house paint. In particular, if the firm is liable even when it is not the cause of losses, it will have an excessive incentive to spend on care. In that case, as its pollution-associated liability would be $180,000, the firm would be willing to spend up to $180,000 on the smoke scrubber to eliminate pollution and avoid liability for the costs of repainting homes.

To illustrate that incentives to engage in activities may be inappropriate if parties are liable even when they are not the cause of losses, suppose that the firm in the example is unable to avert the pollution by taking care (perhaps smoke scrubbers are prohibitively expensive). In this case, the firm can discontinue the activity that generates pollution. Clearly, it should do so if and only if the benefit from the activity is less than $100,000; it is socially desirable for the firm to continue with its activity if the benefit from the activity exceeds $100,000. Again, this is what the firm will do if it is liable for repainting costs if and only if it is their cause. If by continuing its activity it becomes liable for all repainting, that is for $180,000, the firm may be undesirably discouraged from continuing with its activity (that would be so if the benefit from the activity is between $100,000 and $180,000).

The basic function of the causation requirement under strict liability, in other words, is that it furnishes socially appropriate incentives to reduce the risk of harm and to moderate the level of activity by imposing liability equal only to the *increase* in social costs due to a party's actions.[97]

11. 3 Negligence rule. Under the negligence rule, we know from chapter 2 that parties will generally be led to exercise optimal care if due care is set at this level, but the analysis did not take causation into explicit account. If parties do not escape liability for negligence, will they have inappropriate incentives to take care? Is it necessary to allow parties to escape liability in order for their incentives to be correct (as it is under strict liability)? As will be discussed, the answer is that there is no need to allow parties to escape liability for negligence if they do not cause losses, but optimal incentives are maintained even if they do escape liability if they do not cause losses. In other words, basic incentives to take due care are correct whether or not there is a causation requirement. Thus, other factors must be adduced as a warrant for the causation requirement under the negligence rule.

To explain, let us reconsider the example of the polluting firm. Suppose that a smoke

---

[97]This tolling-of-social-costs function of the causal requirement was initially emphasized in an important article, Calabresi 1975, and was amplified in Shavell 1980a in a formal economic treatment of causation and liability.

scrubber costs $30,000, an amount less than the $100,000 of additional harm the scrubber would avert, so that it would be negligent for the firm to fail to install it. Suppose too that if the firm negligently omits to install the scrubber, it will be held liable for harm even if it is not the cause of harm, that is, even if paint peels due to exposure to sun rather than to pollution. This cannot cause the firm to take excessive precautions. All that it does is to increase the firm's incentive to take due care and install the scrubber, for the firm will be threatened with $180,000 of damages for failure to install the scrubber rather than with $100,000. However, a question arises: Could it be that allowing the firm to escape liability for negligence when it is not the cause of peeling paint dilutes unduly the firm's incentive to install the scrubber? The answer is no; the firm will be threatened with $100,000 of damages for the harm it does cause, so it will install the scrubber. Rational actors will always be led to act non-negligently even if they would escape liability when they are not the cause of losses.[98]

Although allowing parties to escape liability if their negligence does not cause losses still leaves sufficient deterrence to induce proper care-taking, an affirmative reason for insisting on causation before imposing liability has not been supplied. One advantage of the causation requirement is that to the extent that there are errors in the negligence determination, the negligence system takes on aspects of strict liability, so that the problem of overdeterrence discussed in section 1.1 would be exacerbated were the scope of liability extended to losses not caused by negligence. A second advantage, probably more important, is that the administrative costs of the liability system are reduced because the volume of cases is lowered owing to the causation requirement.[99] These advantages of restricting liability under the negligence rule to harm caused by the actor may justify the causation requirement.[100]

---

[98]The general proof of this conclusion involves, among other elements, the point that the socially desirable level of care itself implicitly reflects causation; care is socially valuable only to the degree that it can reduce accident losses in circumstances in which losses would otherwise result. This was first demonstrated in Shavell 1980a; see also Shavell 1987a, 105-08, 118-21.

[99]A qualification is that use of the causation requirement for liability means that, in cases that are brought, causation has to be determined (whether the pollution or exposure to sun caused paint to peel), with attendant costs. Consequently, it is possible that use of the causal requirement for liability increases rather than reduces administrative costs.

[100]An additional, somewhat subtle feature of the causation requirement under the negligence rule is worth mentioning. Under the negligence rule, a person's expected liability is often assumed to rise discontinuously with his level of care: Liability is zero if a person is not negligent, yet becomes distinctly positive as soon as his behavior crosses the negligence threshold. For example, if it is negligent to drive at speeds exceeding 50 mph, a person's liability will be zero if he drives at speeds up to 50 mph but suddenly will rise if he drives at 51 mph, for then he will be liable for all accidents that he causes. This jump in expected liability makes the incentive to be non-negligent sharp, which has both socially advantageous and disadvantageous aspects. It may be socially advantageous because it means that parties will have incentives to be non-negligent even if they cannot pay for the entire harm, or even if they will not always be sued for harm; see Cooter 1982 and Shavell 1986. And it may be socially disadvantageous because it means that parties may be led to take excessive care to reduce the risk of mistakenly being found negligent and bearing liability; if people drive at 45 mph they will lower the chance of erroneously being clocked at a speed of over 50 mph; see section 1.1. But, as originally noted by Grady 1983 and Kahan 1989, there may not be a sudden increment in expected liability -- expected liability will rise continuously -- if liability for negligence is properly limited by the causal requirement under discussion. If a person drives at 51 mph, it might be thought that he

11.4 Proximate causation. Even if a party is shown to be a cause of losses, he may still escape liability because he was said not to be the proximate cause of losses, where this term has two major meanings (additional meanings will not be reviewed here).[101] One connotation of proximate cause is that harm came about in a direct or expected way, rather than in an unusual, freakish manner. The legal policy of relieving a party of liability for unusual accidents was discussed in sections 6.2 and 6.3, where no clear justification was apparent.

A second notion of proximate causation concerns coincidence, and is illustrated by two cases. Suppose that a speeding bus happened to be at just the "right" point on a road to be struck by a falling tree. Here, note that the excessive speed of the bus did cause the accident -- had the bus not been speeding it would not have been struck -- but the accident would also have been avoided if the bus had been going even faster. In the other case, a person negligently handed a loaded gun to a child to be used as a plaything and the child dropped the gun on his toe, suffering an injury. In such cases, liability is not found because of lack of proximate cause in the sense that the accidents are said to be coincidental to defendants' behavior, unrelated to the normal risk created by their behavior.

It can be demonstrated that allowing parties to escape liability for accidents like these does not lead to inadequate precautions. Holding a bus company liable when trees fall down on buses will not induce the company to have its buses go more slowly, for the probability of a bus being struck by a falling tree does not depend on the speed of the bus. Likewise, holding a person liable when an object he gives to a child drops on the child's toe will not induce people to remove bullets from guns, for the probability of a gun dropping on the child's toe will not be affected by its being loaded (setting aside the negligible weight of the bullets).[102]

An advantage of permitting defendants to escape liability when accidents are coincidentally caused is that this will lower administrative costs by reducing the scope of liability, unless the cost of deciding about the issue of coincidental causation exceeds the savings from the reduction in the scope of liability. A disadvantage of allowing defendants to escape liability for coincidental accidents is that this means that actors do not bear the full increase in social costs due to their activity (if people did not ride in buses, they might not be struck by falling trees), and the control of levels of activity is an object of, at least, strict liability.

11.5 Uncertainty over causation. In many situations there is uncertainty about causation. For example, it may not be known which manufacturer out of many sold the product (a drug, lead paint) that caused the injury, or whether an injury was caused by the defendant or by background factors (was cancer caused by a firm's pollutant or by unknown environmental or

---

should be liable only for accidents that were caused by going the extra mile per hour beyond 50 mph; the driver should not be liable for accidents that would have occurred had he been traveling at 50 mph or less. However, if the court is unable to tell whether or not an accident would have occurred had a person been driving more slowly (often a plausible assumption), the person driving 51 mph will be liable for any accident that his driving causes, and there thus would be a jump in his expected liability.

[101] Other meanings will not be reviewed here. On proximate causation, see, for example, Dobbs 2000, chapter 10, Hart and Honoré 1985, and Keeton 1963 on Anglo-American law, and Honoré 1983, sections 20, 80-90, and Zweigert and Kötz 1998, 601, 621, on the law in France and Germany.

[102] See Calabresi 1975, Shavell 1980a, and Shavell 1987a, 110-15, 121-23; the latter references formalize the idea of coincidence as illustrated by the two cases mentioned in text.

genetic determinants?).

The law takes two approaches in such situations. The traditional approach is to hold a defendant liable if and only if the probability that the defendant was the cause of losses exceeds 50 percent.[103] This approach may lead either to inadequate or to excessive incentives to reduce risk. Suppose that a firm sells to only 20 percent of the market. Then the likelihood of the firm being the cause of losses from a product-related injury will lie below the 50 percent threshold and it will escape liability for any harm caused by its product. Consequently, the firm will have no liability-related incentive to take precautions. If, however, a firm's market share exceeds 50 percent, the firm will be liable for all harms due to the product it sells -- and for all harms due to the products that others sell -- for it will always be more likely than not the cause of harm. Thus, the firm's liability burden will be socially excessive. These potential problems of inadequate and of excessive incentives may arise under any liability criterion based on a threshold probability of causation; they are not unique to a 50 percent threshold.[104]

The second approach that the legal system has taken (though not often) is to hold defendants liable despite any uncertainty over causation but to impose damages only in proportion to the likelihood of causation.Thus, liability has been imposed according to the share that firms have in the market for a product.[105] Under the proportional liability principle, it is readily shown that incentives to reduce risk are proper. If, for example, a firm has 20 percent of the market, it will pay 20 percent of harm in every case, so that its liability bill will be the same as if it pays for all the harm in the 20 percent of cases it truly causes -- in which case we know that safety incentives will be socially appropriate. That the proportional liability principle engenders optimal incentives (without a need to establish causation in particular cases) is an advantage of the principle relative to the traditional threshold probability criterion.[106]

---

[103]See, for example, Dobbs 2000, 420-22; on civil law countries, see Honoré 1983, sections 201-03.

[104]Essentially this point has been frequently mentioned; see, for example, Tribe 1971, and Landes and Posner 1983, and it is formally developed in Shavell 1985b.

[105]This principle has been applied by some courts in cases involving the drug DES, which was made by many companies but the identity of the company that produced the drug that caused harm in particular instances often could not be determined. The proportional liability principle has so far not been widely used. See, for example, Dobbs 2000, 430-32.

[106]See Rosenberg 1984 and Shavell 1985b, 1987a, 115-18, 123-26. It should be noted that if products about

Yet a disadvantage of the proportional approach is that it could lead to a substantial increase in the volume, complexity, and cost of litigation, for under the proportional approach any party for whom the probability of having caused a loss is positive can be sued and may have to pay damages. Hence, the proportional approach should be employed only if the incentive advantage of so doing is sufficiently strong.

---

which there is uncertainty as to causation differ in the risks they create, this must be recognized in the application of proportional liability. For example, if two products each command half the market but the first creates twice the risk of the second, then the first product should bear two thirds of damages, not one-half.

Chapter 5

Liability, Risk-bearing, and Insurance

The accident problem involves not only the goal of appropriately reducing the risks of accidents, but also a second objective: allocating and spreading the risk of losses from accidents that do occur, so that those who are risk averse do not bear them, in whole or in part. Insurance -- both accident insurance for victims and liability insurance for injurers -- provides a method of allocating and spreading risks of loss.

Liability insurance is of particular interest for two reasons. First, liability insurance changes the way in which the liability system functions to reduce accident risk, for to the degree that an injurer owns insurance coverage, he will not have to pay damages if found liable. The incentives to reduce risk thus have to do with the extent of insurance coverage and with other terms of liability insurance policies. Second, and related, the question arises whether, or in what sense, liability insurance is socially desirable if it might interfere with liability-related incentives to reduce risk, and thus whether liability insurance should be regulated. (As will be described below, serious doubts about the social wisdom of liability insurance have been raised, and its sale is subject to some restrictions today, although it is widely owned.)

I will begin by reconsidering the accident problem in the light of risk-aversion. Specifically, I will discuss the socially ideal solution to the accident problem, what occurs in the absence of liability, what occurs in presence of liability when insurance is not available, and finally what occurs in the presence of both liability and insurance. This step-by-step method of analysis will give us an understanding of the separate value of both liability and insurance in reducing accident risks and in allocating accident risks.

I will then address two important extensions to the foregoing analysis: nonpecuniary losses, and the judgment-proof problem.

1. Risk Aversion and the Socially Ideal Solution to the Accident Problem

1.1 Risk aversion. In this book *risk aversion* should be understood as a term of art, describing an attitude of dislike of pure financial risk. A risk-averse person would pay to avoid a risk, such as one involving a 50 per cent chance of losing $1,000 and a 50 per cent chance of winning $1,000.[107] A risk-averse person would also purchase insurance against risk if the insurance premium is actuarially fair.[108] A person will be risk averse if the marginal utility of

---

[107]This risky situation involves an expected return of zero: $50\% \times \$1,000 - 50\% \times \$1,000 = 0$. But a risk averse person would also pay to avoid a risky situation with a positive expected return, such as one in which there is a 51 per cent chance of winning $1,000 and a 49 per cent chance of losing that amount. At some probability of winning, however, even a risk-averse person would become just willing to participate in the risky situation. This probability reflects the individual's degree of risk aversion (the higher the probability needed to induce him to participate, the higher the degree of risk aversion).

[108]An insurance premium is said to be actuarially fair if it equals the expected cost of coverage to the insurer. For instance, if a policy pays coverage of $1,000 with probability 10 per cent, the fair premium is $100, for the expected cost

money to him declines as his wealth increases; the reason is essentially that, for such a person, losing an amount of money will reduce his utility more than gaining the same amount of money will increase his utility.

Risk-aversion is most relevant in situations in which losses would be large in relation to a person's assets and thus would impinge substantially on his utility. Individuals are typically viewed as risk-averse actors in relation to serious accidents, as these would be likely to cause losses that are significant in relation to their assets. If, however, losses would be modest relative to a person's assets, he would be likely to display a roughly risk-neutral attitude toward them. This would be the case of small accidents that individuals suffer. Also, firms might be usually be considered as risk neutral actors in relation to many accidents, for these would cause losses that are small in relation to their assets.[109] Moreover, firms are sometimes treated as risk-neutral if they are owned by well-diversified shareholders, for being well-diversified, the shareholders should not be concerned about the risk borne by a particular firm. However, to the degree that the managers of a firm are risk-averse and have the freedom to make decisions, the behavior of the firm may reflect risk-aversion.

1.2 Socially ideal solution. Social welfare will be taken here to be a function of individuals' *expected utilities*.[110] Therefore, social welfare will depend not only on the factors emphasized in previous chapters -- positively on the benefits parties obtain from their activities, and negatively on costs of care and on accident losses -- but social welfare will depend as well on whether risk-averse parties bear risk. It follows that, under the socially ideal solution to the accident problem, two things will be true: Not only will parties make decisions about engaging in activities and about their exercise of care in the way that was described as optimal in chapter 2, but also, risk-averse parties will not bear risks, which is to say, their risks will be perfectly spread through insurance arrangements or will be shifted to risk-neutral parties. It is important to emphasize that the risk-averse parties who ideally ought not to bear risk may be victims *or* injurers. Injurers will bear risk if they face the risk of liability; thus, ideally, risk-averse injurers should not bear this liability risk.

2. The Accident Problem in the Absence of Liability and Insurance

Assume here that there is no liability system and also no insurance system. Then because injurers will not be liable for accident losses, they generally will not reduce risk appropriately. That is, they will tend to engage in risky activities to an excessive extent and will have no motive to take care. Still, injurers will bear no risk; this aspect of the outcome is socially desirable if injurers are risk-averse.

---

of coverage is $100.

[109]If a firm faces the risk that it would cause harms that are small for each victim but the harms would simultaneously affect many victims (as might be true if there is a defect in the design or manufacture of a product), then the risk for the firm could be substantial, so that it might be appropriate to regard the firm as risk-averse.

[110]The expected utility of an individual is the sum over all possible outcomes of the probability of each outcome multiplied by its utility.

Because victims will not be able to obtain judgments from injurers in the absence of liability, they will be left bearing risk. This bearing of risk is socially undesirable if victims are risk-averse. Thus, the outcome is undesirable not only because risks are not reduced by injurers, but also because the risks that exist may be borne by uninsured risk-averse parties.

3. The Accident Problem Given Liability Alone

Assume now that injurers are subject to liability, but that insurance is again unavailable; thus, injurers do not possess liability insurance coverage and victims do not hold accident coverage. In this situation, the outcome is, in essence, that injurers will be led to reduce risk due to the effect of liability, but the allocation of risk will depend on whether the form of liability is strict or the negligence rule.

In particular, under strict liability injurers will have a motive to reduce risk and victims will, by definition, be compensated for any losses they sustain; it is injurers who will bear risk. If injurers are risk-neutral, their bearing of risk will not matter, and the outcome will be socially ideal.

But if strictly liable injurers are risk averse, the outcome will not be socially ideal because injurers will bear risk. Moreover, they may be led to exercise excessive care to avoid liability (consider how cautiously risk-averse and uninsured individuals would drive if subject to strict liability). In addition, for these reasons injurers may be undesirably discouraged from engaging in an activity. One way of alleviating these problems of excessive care and too low a level of activity under strict liability is to reduce damages; indeed, it can be shown to be beneficial for damages to be less than harm for this reason. In other words, if injurers are risk averse, it is not socially desirable to "internalize" fully the harm they do.[111]

The situation is quite different under the negligence rule, because injurers will not bear risk provided that they take due care (and that the courts accurately assess their level of care), which they will decide to do. Hence there will be no particular problems respecting injurers when they are risk averse; they will not be led to take excessive care nor be undesirably discouraged from engaging in an activity. Victims, on the other hand, will bear their losses (presuming that injurers are not mistakenly found negligent). As a consequence, social welfare will be less than optimal if victims are risk averse and are not insured.

The foregoing points thus introduce a new element into the comparison of strict liability and the negligence rule. Under strict liability, risk will be borne by injurers, whereas under the negligence rule risk will be borne largely by victims. In the absence of insurance, therefore, the relative appeal of strict liability will be enhanced when injurers are risk neutral or, more generally, when they are less risk averse than victims, and the relative appeal of the negligence rule will be enhanced when victims are less risk averse than injurers.

---

[111]See Proposition 2 in Shavell 1982a.

4. The Accident Problem Given Liability and Insurance

4.1 Insurance. Now let us assume that insurance is available and sells at actuarially fair rates. One form of insurance is, as mentioned, accident insurance for victims. If victims are risk-averse, they will buy complete coverage against any risk that they bear. Thus, for example, if injurers take due care under the negligence rule, so that victims bear the risk of accidents, victims will fully insure against these risks if they are risk averse.

The other form of insurance that we will consider is insurance for injurers against liability. If injurers are risk-averse, they will wish to purchase such liability insurance. However, the insurance purchase decision, and the nature of insurance policies, becomes somewhat complicated because ownership of insurance may itself change the incentives of injurers to prevent accidents, and thus the likelihood of liability. This will be discussed below.[112]

4.2 Strict liability. As stated in section 3, when liability is strict, victims will be implicitly insured by the legal system, and so will not bear risk, but injurers bear risk. Therefore, risk-averse injurers will wish to purchase liability insurance. In order to describe their insurance purchases, we will consider two standard cases about liability insurers: one in which the insurers can observe the level of care exercised by insureds and are able to link premiums to this factor, and another in which the insurers cannot observe the level of care and base premiums on it. In both cases, we will conclude that liability insurance is socially desirable.

Liability insurers can observe care. Suppose first that insurers can observe the level of care and thus can lower the premium to reflect the risk reduction that care engenders. An insurer might, for instance, inspect an insured injurer's building to see whether or not he purchased fire extinguishers to decrease the risk of liability for fires, and lower the premium if extinguishers were purchased. In such a setting, two points are true. First, insureds will purchase full coverage to protect themselves completely against risk. Second, insureds will be led by fair premium reductions to take optimal care. Suppose that an injurer purchases full coverage against a $1,000,000 liability for harm to others from fire and that extinguishers would lower fire risk from 5 per cent to 1 per cent, so would lower expected liability by $40,000. Then premiums for full coverage will be reduced by $40,000 if the extinguishers are purchased, implying that the insured will buy the extinguishers as long as their cost is less than $40,000, which is to say, the injurer's decision to take care will be optimal.

It should be noted from what has been said that the outcome will be socially optimal. Risk-averse injurers will be fully protected against risk and will be induced by premium reductions to take optimal care. And victims will be protected against risk by definition of strict liability.

Because the outcome is socially optimal, liability insurance is socially desirable. To put the point differently, were liability insurance not present, or forbidden, injurers would be worse off because they would bear risk (and might take excessive care or engage too little in desirable activities), whereas victims would be just as well off.

Liability insurers cannot observe care. Now suppose that insurers cannot observe the

---

[112]See Shavell 1987a, chapters 8, 10, for a presentation of the elements of the theory of insurance that are relevant to this chapter.

level of care and therefore do not charge a premium that depends directly on the level of care an injurer exercises. We might imagine, for instance, that in the example of fire risk, the level of care corresponds to the caution with which flammables are handled, something that an insurer would have a more difficult time checking than the presence or absence of fire extinguishers. In a case like this, injurers will usually purchase only partial coverage and the level of care will tend to be less than optimal.

To explain, observe first that because the assumption is that insurers cannot observe and thus cannot penalize injurers for not being careful with flammables, injurers will have no premium-related incentive to take extra precautions (like training workers in the handling of flammables). Consequently, if injurers have full coverage, the risk of a $1,000,000 fire liability would be 5 per cent, and premiums would thus have to be $50,000. If, however, coverage is partial, then injurers will have some incentive to take care because they will suffer some uncompensated losses if held liable. For example, suppose that they have $600,000 of coverage, exposing them to a $400,000 risk. Therefore, a precaution that lowers risk from 5 per cent to 1 per cent would have value to them of at least $4\% \times \$400,000$ or $16,000. This factor in turn may lead injurers to prefer partial coverage. Suppose that the precaution costs $10,000. Then if coverage is $600,000, injurers will be led by their exposure to risk to take the precaution, for spending the $10,000 will save them $16,000 in expected losses that they will bear. And because the risk for injurers who buy partial coverage of $600,000 will thus be only 1 per cent, the premium they will be charged will be $6,000.[113] Injurers might well prefer 60 per cent coverage that costs them only $6,000 to full coverage that costs them so much more, $50,000. Such policies with less than complete coverage, however, do not expose injurers to enough risk to induce them always to take optimal precautions.

The outcome in the situation under consideration is thus not socially ideal for two reasons: Injurers tend to be only partly protected against risk, and their level of care tends to be less than optimal.[114] However, victims are protected against risk by definition of strict liability.

Nevertheless, liability insurance is socially desirable. (That the outcome is not socially ideal does not mean that liability insurance is undesirable.) One way of seeing that liability insurance is socially desirable is to notice that it obviously raises the well-being of injurers -- after all, they choose to purchase liability insurance -- and does not affect the well-being of victims -- they are fully compensated for accidents whenever they occur. Thus, the presence of liability insurance raises social welfare.[115] Conversely, were liability insurance prohibited,

---

[113]To amplify, insurers will charge $6,000 because injurers who purchase coverage of $600,000 will suffer losses and make claims with only a 1 per cent probability. Insurers do not need to observe in a direct way that the injurers are led to take care; the insurers need only note the statistical fact that injurers who buy the $600,000 partial coverage policy make claims with a 1 per cent probability.

[114]In the example, care is discrete in nature: it is either taken or not. Therefore, it turns out that if care is optimal to take, either there exists a partial coverage policy that insureds prefer that induces care to be exercised, or else there does not, in which case a full coverage policy is purchased. In other words, either coverage is partial and care is optimal, or coverage is full and care is zero. But in a model with care continuously variable, it is usually true that coverage is partial and that the level of care is less than optimal.

[115]This argument is not really a proof, because it presumes that the level of damages equals harm. In fact, that the level of

injurers would be exposed to risk and made worse off (and perhaps would not engage in desirable activities and/or would take excessive care), whereas victims would not be helped.

4.3 Negligence rule. Under this rule, we will argue that injurers will tend to take due care even though they can buy liability insurance, and that liability insurance will tend to be socially desirable. Because injurers tend to take due care, victims will bear the risk of losses and will purchase accident insurance if they are risk averse.

Consider initially a perfectly functioning negligence system. In such a regime, the basic logic from chapter 2 about behavior under the negligence rule implies that injurers will be led to take due care, assuming that they are not insured. In our example, suppose that a $10,000 precaution reduces the $1,000,000 fire risk from 5 per cent to 1 per cent, so that the precaution is required to avoid negligence. An injurer will take the precaution, because that will free him of the liability he would otherwise bear, a 5 per cent chance of a $1,000,000 liability. (More precisely, it is clear that the injurer will be led to take the precaution if he is risk neutral, and he will want to take it more strongly if he is risk averse, which is the assumption.)

Although an injurer will wish to take due care if he is *not* insured, might he want to purchase liability coverage in order to act negligently? The answer is no -- because the insurance would cost him too much. If an injurer were to purchase an insurance policy that covered him for negligently caused harm, he would decide not to take the precaution, that is, to act negligently, and thus would cause harm of $1,000,000 with probability 5 per cent. Hence, the premium for the insurance policy would have to be $50,000. Clearly, when faced with the choice of paying $50,000 for the insurance policy in question, or instead spending $10,000 on the precaution so as not to be negligent, the injurer would do the latter.

In the basic model of a perfectly functioning negligence rule, therefore, the outcome is socially ideal. Injurers are led to take due care, and, being non-negligent, do not bear risk. Victims purchase accident insurance if they are risk averse so they do not bear risk if they want to avoid it. Liability insurance is not undesirable, because it does not interfere with the deterrent of the negligence rule; indeed, liability insurance is essentially moot, as it is not purchased.

Next consider briefly the realistic situation in which the negligence rule does not function perfectly and there is a risk of findings of negligence due to errors in the negligence determination or to inability of injurers to control perfectly their levels of care (see section 1 of chapter 4). In this case, the main difference to note is that injurers might be found negligent even if they try to take due care. Thus risk-averse injurers will decide to purchase liability insurance, and the type of policy risk-averse injurers will purchase will protect them primarily against being found negligent through some type of error or lapse. The policy will not protect injurers so broadly as to induce them definitely to act negligently because, as explained above, a policy that induced injurers definitely and intentionally to act negligently would not be purchased because the premium would be too high. Therefore, the liability insurance coverage tends not to compromise deterrence, and is socially desirable because it protects injurers against risk. Moreover, its ownership reduces the problem of excessive care caused by uncertainty in the negligence determination (see again section 1 of chapter 4).

---

damages equal harm can also be shown to be optimal; see Shavell 1982a.

4.4 Summary. Three points about liability and insurance summarize the analysis to this point. First, because liability insurers pay for some or all of the losses for which injurers are found liable, the manner in which liability rules alter injurers' behavior is to a significant degree indirect, being associated with the terms of their liability insurance policies (notably, the connection between premiums or the payment of claims and injurer behavior, and the level of coverage).[116]

Second, the availability of liability insurance is socially desirable. The particular arguments demonstrating this result depended on the form of liability and insurers' information about insureds' behavior. The arguments were, roughly, based on the following considerations. The availability of liability insurance increases the welfare of risk-averse injurers because it protects them from risk and ameliorates the problems that they would otherwise take excessive care or be discouraged from engaging in desirable activities. Moreover, the availability of liability insurance does not necessarily dilute injurers' incentives to reduce risk, and where it does do that, the dilution of incentives will be moderate, for policies that would substantially increase risks would be so expensive that they would not be attractive for purchase.

This conclusion about the social desirability of liability insurance is not to deny the possibility that liability insurance might be socially disadvantageous, and I will discuss below circumstances in which it should be regulated (see section 7). But the conclusion does mean that thinking about the social desirability of liability insurance should proceed from the understanding that, as in the basic model of liability studied here, the insurance is socially desirable.

Third, the availability of accident and liability insurance limits the importance of the allocation of risk as a factor to be considered in evaluating liability rules. For example, the fact that in some areas of accidents the typical injurers might be large, essentially risk-neutral firms and the victims risk-averse individuals will not constitute an argument in favor of imposing liability to the extent that the individuals are insured against their losses.

4.5 Liability and insurance in reality. The importance of insurance to the liability system is very great, so that the theory of this section, rather than that of section 3, is the most relevant for understanding the liability system. In particular, accident insurance is widely held.[117] Liability insurance is also owned by many individuals, and an indication of its salience is that over 90 per cent of all payments made to tort victims are paid by liability insurers.[118] Although

---

[116]Other possible terms of coverage, though not discussed above, are the link between loss history and premiums or the future right to insure.

[117]Notably, approximately 86 per cent of the population possesses health insurance benefits; see U.S. Census Burean 2001, 2. Also, approximately 85 per cent of husband-wife families with children possess life insurance on at least one family member; see ACLI Fact Book 1999, 1. Additionally, in 2000 about 88 per cent of the adult population was covered by the Social Security system; see Statistical Abstract of the United States, 2001, 13, 345. Moreover, over half the workforce possesses some form of private disability coverage; see U.S Department of Labor 2001, 1.

[118]See, for example, O'Connell et al. 1994, App. A, from which it is evident that total liability payments made in 1990 were $65.199 billion, of which $60.981 billion were made by liability insurers; thus about 93.5 per cent of tort liability payments were made by liability insurers. See also Tillinghast-Towers Perrin 2002, App. 4, which reports

commentators occasionally raise the issue that liability insurance might undermine deterrence, and there are some restrictions on its sale (see section 7 below), it is generally legal. Historically, however, the sale of liability insurance was resisted, and in some countries it was not permitted until the early twentieth century.[119] Perhaps the most notable instance of antagonism against liability insurance was the complete ban on its use in the former Soviet Union.[120]

   Note on the literature. The functioning of the liability system together with insurance is first formally studied in Shavell (1982a, 1987a), where the social desirability of liability insurance is proved; this chapter largely follows my treatment there. However, an early and insightful informal examination of the subject is contained in Calabresi (1970).

5. The Purpose of Liability

   5.1 Compensation of victims is the traditional conception of the purpose of accident liability. The great majority of legal scholars, lawyers and judges, and, probably, citizens appear to assume that providing victims with fair compensation for harm is the primary purpose of accident liability. A representative statement is that of Prosser: "There remains a body of law which is directed toward the compensation of individuals. . . . This is the purpose of the law of torts."[121] That the object of accident law should be so viewed is not surprising, for the goal of the plaintiff in an action is generally to be compensated and that of the defendant is to avoid paying the plaintiff; the two do not usually consider deterrence to be nearly as important, if it is an issue. Moreover, the classical and intuitively appealing notion of corrective justice, that a wrongdoer should compensate his victim, comports with the view of compensation as the purpose of liability (at least under the negligence rule). This is not to say that deterrence is never seen as an additional purpose of accident law; one sometimes reads statements by commentators and judges to that effect. However, rarely does one encounter the belief that the main purpose of accident law is deterrence and not compensation.

   5.2 Reduction of risk through deterrence of harm is the true purpose of liability today, but compensation and avoidance of strife were also important historically. In contrast to the traditional view, and as the discussion of this chapter should make clear, compensation cannot be said to be a primary purpose of accident liability if, as is the case, accident insurance is largely available to victims (and could be provided to them through social insurance if need be). In other words, in the absence of the liability system, compensation of

that about 98 per cent of personal tort costs were paid by liability insurers in the period 1973-2000.

   [119]As Tunc writes in his survey of tort law worldwide, "At the beginning of the nineteenth century, liability insurance would have been unthinkable. It would have been considered as immoral." He goes on to mention, among others, French and Scandinavian resistance to the sale of liability insurance. Tunc 1983, 50-51.

   [120]On the Soviet ban on coverage, see generally Rudden 1966; see also Tunc 1983, 51-52.

   [121]Keeton, Dobbs et al. 1984, 5. See also, for example, Fleming 1998, chapter 1. Fleming is quite direct in stating that deterrence is not a primary purpose of accident law; he says at p. 10 that "in the core area of tort accidents . . . It is being increasingly realized that human failures in a machine age exact a large and fairly regular toll . . . which is not significantly reducible. . . through the operation of tort law."

victims would probably be about as well accomplished through private and social accident insurance as it is today. The main difference that the liability system can make to outcomes is the creation of incentives toward safety. (Further, as explained in section 4, we know that the liability system can generate these incentives despite the existence of liability insurance.) Hence, if the liability system has a real purpose today, it must lie in the creation of incentives to reduce risk.[122]

It is worth noting, however, that the function of accident liability was different in the past. Before the development of insurance markets in the latter part of the nineteenth century, liability furnished victims a source of compensation that would not otherwise ordinarily have been forthcoming. Tort liability law thus served to an important degree the dual purposes of compensation and deterrence. Moreover, in early times before criminal law and tort law had emerged as separate branches of law, a significant additional purpose of the making of money payments for harm was the maintenance of social order. Without the system of money payments, private vengeance would often have followed the doing of harm.[123]

5.3 Is use of the liability system justified? It has been explained that the main route through which the liability system benefits society is by reducing accident risks, so that this, if anything, should be the warrant for use of the liability system. However, the deterrent benefit of the liability system is not necessarily sufficient to outweigh the costs of the system and thereby to justify its use. As I will discuss in the next chapter, the costs of the liability system are great. Moreover, safety regulation is available as an alternative means of reducing risk. Therefore, the question of the net social desirability of the liability system is a serious one.

6. Extension: Nonpecuniary Losses

6.1 Nonpecuniary losses and insurance. Let us consider the relationship between nonpecuniary losses and accident insurance before re-examining the functioning of the liability system.[124] To understand insurance for nonpecuniary losses, observe that suffering a nonpecuniary loss often will not alter an individual's need for money or, more exactly, the utility he would derive from receiving additional money. If, for example, an irreplaceable family portrait with great sentimental value is destroyed, there is no obvious reason to believe that the owner's need for money will increase, however much he regrets the loss; the utility he would obtain from having more money to spend would be whatever it was beforehand. Similarly, if a

---

[122]To avoid confusion, let me note that I am assuming that the "purpose" of the liability system means the difference to outcomes that the liability system actually makes. This definition of the purpose of the liability system -- as opposed to what people *say* its purpose is -- is both both natural and a guide to policy, for policy must be based on the true effects of the liability system on outcomes, not the supposed effects.

[123]On the point that the making of money payments for harm prevented feuds, when there was no criminal law, see, for example, Berman 1983, 55, and Pollock and Maitland 1911, 2: chapter 8, section 1. When a criminal system exists, this system exacts punishment for many types of act (like murder and rape) that would give rise to feuds, so that accident law is not as much needed to avert feuds.

[124]Arrow 1974, Cook and Graham 1977, and Zeckhauser 1973 first developed the theory of insurance for nonpecuniary losses. For relevant empirical evidence, see Viscusi and Evans 1990.

person loses a small toe in an accident, then aside from requiring some money for medical treatment, he might well place the same value on having additional money as he had prior to losing his toe.

In some cases, though, events with adverse nonpecuniary consequences will result in a person attaching a higher value to money. An individual who is crippled by an accident may value money more, even after being compensated for medical expenses and forgone income, because of a need to obtain household help, special transportation services, and the like.

It is also possible that suffering a nonpecuniary loss will lower the utility of money to an individual. The individual who is crippled by an accident could turn out to value money less because venturing forth to spend is less pleasurable and more difficult. Perhaps the most important example of a nonpecuniary loss that results in a lower value of money involves death. The value of money in that contingency is, in effect, the value to a person of knowing that his survivors will receive a bequest, and this will often be less than the value to the person of having money while he is alive.

The amount of insurance coverage against nonpecuniary losses that an individual will in principle wish to purchase will depend on whether such losses will affect the utility he would derive from receiving additional money. The reason is that the purchase of insurance is in essence the giving up of money today for the receipt of money in a contingency, and this will make sense only if money is more valuable in that contingency.

Now if nonpecuniary losses will not result in a change in a person's valuation of money, then the best insurance policy for him, the policy under which his expected utility would be maximized, will not cover him for the nonpecuniary element of his losses. His coverage will be restricted to pecuniary losses, if any. Thus, a person might not insure against the loss of his family portrait and might limit coverage against loss of a toe to medical expenses. Notice that this implies that, under the optimal insurance policy, the coverage a person will receive will not make him whole in utility terms (even though this may be possible). In other words, the notion that optimal insurance for a loss will restore his utility to its level before the loss is incorrect.

If, however, nonpecuniary losses will raise the value of money to a person, then under the optimal insurance policy coverage will exceed his pecuniary loss. Thus a person might purchase greater coverage against the possibility of being crippled than an amount equal only to the costs of medical treatment and forgone earnings. It is unlikely, however, that he will purchase coverage sufficient to make him whole (if this is possible).

If the value a person will place on money will decrease as a result of a nonpecuniary loss, expected utility maximizing insurance coverage will be less than his pecuniary losses. A person who has little desire to leave a bequest will rationally purchase little or no life insurance, despite the possibility that the earnings forgone by his death will be large.

6.2 Actual insurance coverage against nonpecuniary losses. It seems that actual insurance coverage is intended mainly to remedy pecuniary needs created by losses, not to compensate for the disutility due to losses, suggesting that nonpecuniary losses tend not to raise the value of money to individuals or are small. In particular, insurance coverage against loss of property does not ordinarily seem to reflect its sentimental value, only its market value or replacement cost. Coverage against personal injury usually approximates only direct medical

expenses and forgone earnings. Insurance against death is ordinarily bounded by lost earnings; if a person (such as an unmarried or elderly individual) has no dependents, he normally possesses little or no insurance coverage; parents do not often carry significant coverage on the lives of their children.

6.3 Nonpecuniary losses and the socially ideal solution to the accident problem. As just emphasized, the amount of insurance an individual would wish to purchase against nonpecuniary losses -- and therefore the amount of money he will receive under the socially ideal solution to the accident problem -- will be based on the value he will place on money if he suffers nonpecuniary losses; the amount he will receive will not generally make him whole. By contrast, and as was noted in section 8 of chapter 4, the socially optimal level of care taken by injurers (and their level of activity) will fully reflect the nonpecuniary elements of accident losses as well as the pecuniary. Thus, for instance, it will be best that injurers take substantial care to reduce the risk of accidentally killing children even though their deaths may not impose an economic burden on their parents and consequently, in the ideal, may not call for the parents to receive significant amounts in compensation. It might be, say, that for injurers to be led to take appropriate care, they should pay $2,000,000 for the death of a child, but that the parents' optimal insurance coverage is only $10,000 for funeral expenses.

6.4 Socially ideal solution cannot be achieved under the liability system. Because injurers must exercise a degree of care reflecting both nonpecuniary and pecuniary components of victims' losses under the ideal solution to the accident problem, the magnitude of payments that injurers make under liability rules has to reflect both these components of losses for injurers to be led to take optimal care. But if injurers' damage payments are this high, then the amount victims receive will exceed optimal compensation, which may well usually approximate only pecuniary losses. If parents receive $2,000,000 for the wrongful death of a child, this will exceed the $10,000 optimal compensatory amount. Yet if injurers' damage payments equal only the $10,000 optimal compensatory amount, injurers' incentives to take care will be inadequate. Thus the socially ideal outcome cannot be achieved under the liability system.[125] The damage payments will inevitably result in a compromise between awarding victims correctly and creating appropriate incentives for injurers to reduce risk.[126]

6.5 The case for fines as a supplement to liability. An improvement over the situation with the liability system may be achieved by a regime in which liability is supplemented by fines collected by the state.[127] With the use of fines, the total amount that injurers are made to pay can

---

[125]This statement is correct under strict liability, but under the negligence rule there is a qualification to it. Under a perfectly functioning negligence rule, the socially ideal outcome can be achieved, because under that rule injurers will always take due care and never be found negligent if damages are sufficiently high (include nonpecuniary losses). Hence, victims will bear their losses and can and will optimally insure for amounts less than what injurers would have to pay were they liable. But if, as is realistic, findings of negligence occur, the optimal outcome cannot be achieved because victims will in fact sometimes receive awards.

[126]This section is, of course, premised on the assumption that victims of harm are not buyers of a product sold by injurers. If victims are buyers, then an ideal outcome can clearly be achieved, as long as liability does not exceed the optimal compensatory amount. On the latter issue, see, for example, Rubin 1993.

[127]Spence 1977 first demonstrated the desirability of employing fines in addition to liability as a result of nonpecuniary

be raised to the point that their incentives to reduce risk are appropriate, while at the same time liability can be held to the lower level equal to optimal compensatory awards. Thus, in the example, under the contemplated regime injurers would pay fines of $1,990,000 reflecting nonpecuniary losses, whereas victims would receive in liability awards payments of $10,000 reflecting only their otherwise uncompensated pecuniary losses.

One way to understand why individuals may find a regime with supplemental fines advantageous is to recognize that their taxes can be lowered under it, because the state may use fine revenues to replace tax revenues. Specifically, individuals will find a regime with fines advantageous if they would prefer a savings in taxes to collecting higher liability judgments, or, equivalently, if they would not be willing to pay the insurance premium necessary to purchase coverage in the amount of the fine. Parents should thus find advantageous a regime with fines for the wrongful death of their children and with correspondingly lower taxes if the parents do not choose to insure their children's lives.[128]

Several comments should be added about supplemental fines. First, it would ordinarily be best for these fines to be insurable. The general argument made in section 4 that it is desirable to allow risk-averse injurers to purchase coverage against liability can be employed to demonstrate this result.[129]

Second, to calculate the magnitude of supplemental fines for nonpecuniary losses, one can employ extrapolations from the amount individuals would be willing to pay for a small reduction in the probability of suffering nonpecuniary losses. Suppose, for instance, that an individual would be willing to pay $1,000 for a 1 percent reduction in the likelihood of losing his arm. Then the optimal fine for causing the loss of his arm would be approximately $100,000. Information about persons' willingness to pay for reductions in risk could in theory be obtained by survey or, in some cases, perhaps, by attributing wage differences to differences in risks of accidents.[130]

Third, the argument favoring supplemental fines applies more broadly than here, to any situation in which the amount that injurers should pay to be properly deterred exceeds the

---

losses.

[128]To amplify, suppose that the likelihood of death of a child is .1 per cent. Suppose too that a parent would not want to buy an insurance policy for $2,000,000 if a child died but rather would limit coverage to $10,000. Hence, the supposition is that the parent is worse off with coverage of $2,000,000 and a premium of $2,000 (that is, .1% × $2,000,000) than with $10,000 coverage and a premium of $10 (that is, .1% × $10,000). In other words, increasing coverage from $10,000 to $2,000,000 is not worth the $1,990 increase in premiums that would be required. But this statement implies that the parent is better off under a liability system that gives the parent $10,000 in damages and imposes a fine of $1,990,000 than under a liability system that gives the parent $2,000,000 in damages; for under the fine system, the state's expected revenues increase by $1,990, lowering the parent's taxes by that amount.

[129]It should be mentioned, however, that if fines are not employed, then in principle it could be advantageous to set liability at a level approximating optimal compensation and to limit purchase of liability insurance, so as to induce injurers to take more care.

[130]On the latter, see Moore and Viscusi 1990 and Viscusi 1983. See also Danzon 1984, which discusses how such data can be used to calculate supplemental fines.

optimal compensatory amount for victims. Notably, it was emphasized that when liable injurers will not always be identified as responsible for harm done, the amount they pay if they are identified and sued must be raised so that their incentives to reduce risk will be maintained at the correct level. When, for instance, the likelihood is 50 percent that a liable injurer will be successfully sued, the amount he pays if sued must be on the order of twice the victim's losses. Optimal payments by injurers may therefore exceed optimal compensatory awards by a substantial factor, so fines would be desirable. Another reason why optimal payments by injurers may be greater than optimal awards concerns taxes that would have been paid on income forgone by accident victims. For injurers' incentives to reduce risk to be proper, they must pay an amount based on before-tax income forgone by victims.Yet the amount of money that victims will in fact lose, and thus the amount that will constitute the optimal compensatory award, is after-tax forgone income. An additional reason why optimal payments by injurers may exceed optimal compensatory awards is that victims may receive insurance benefits or gifts. In this case, optimal compensatory awards will equal only the shortfall between victims' receipts and their losses, but injurers must pay victims' entire losses to be adequately deterred.

7. Extension: The Judgment-Proof Problem

7.1 The judgment-proof problem and insurance. The possibility that injurers do not have sufficient assets to pay for harm has implications both for the purchase of accident insurance by victims and for the purchase of liability insurance by injurers.

First, the incentive of victims to purchase accident coverage is increased due to the judgment-proof problem, for it means that there is now a risk that a victim who ought to be able to collect from a liable injurer is unable to do so.

Second, the motive of injurers to purchase liability insurance is diminished because of the possibility that they would be judgment-proof.[131] The reason is that insuring against liability that one would not otherwise fully bear, because one's assets would be exhausted, is in a sense a waste for a party. An injurer with assets of $20,000 who faces a 10 per cent risk of liability of $100,000 would have to spend $10,000 on premiums for full coverage, 80 per cent of which would be attributable to coverage of the $80,000 that he could not pay in the absence of liability insurance coverage. Consequently, the individual might well decide against buying full liability insurance coverage even though he is risk averse. In general, a risk-averse party might rationally decide to purchase less than complete coverage, or no coverage at all; his purchase decision will depend on what his assets are in relation to the potential liabilities, their likelihood, and his degree of risk aversion.

7.2 Problems with the functioning of the liability and insurance system. It was emphasized before, in section 3 of chapter 4, that the judgment-proof problem dilutes injurers' incentives to reduce risk because they do not have the capacity to pay for the harm they might cause. The fundamental nature of this problem is not altered in the present context, but there are

---

[131]Calabresi 1970, 58, observed that the motive to purchase insurance will be diminished when parties' assets are less than the losses for which they may be held responsible; Keeton and Kwerel 1984 and Huberman, Mayers, and Smith 1983 first investigated the point formally.

two differences in its character worth noting. First, to the extent that injurers are risk-averse and do not own liability insurance coverage -- a real possibility in view of their reduced incentives to purchase such coverage as just discussed -- the problem of insurance-related dulling of incentives to lower risk will be less severe than before; moreover, a risk-averse injurer will tend to take more care to prevent liability than a risk-neutral injurer. Second, to the extent that injurers own liability insurance, the problem of the dulling of their incentives to reduce risk might be exacerbated. This could be so if liability insurers cannot observe injurers' levels of care and link premiums to those levels. However, if liability insurers can observe levels of care, the conclusion would be different.

The judgment-proof possibility not only lowers incentives to reduce risk, it also creates problems in respect to the bearing of risk. In particular, because injurers have a diminished motive to purchase liability insurance coverage, and may buy none at all, they are left bearing risk, which lowers social welfare. However, victims should not bear risk as a result of the judgment-proof problem because they will rationally purchase accident insurance coverage against not being compensated by liable injurers.

7.3 Regulation of liability insurance. The problems that the judgmentBproof issue creates leads to the possibility that it would be desirable to regulate liability insurance, and in either of two contrasting ways -- to forbid its purchase, or to require its purchase.[132] We will discuss the circumstances in which each type of regulation may be socially advantageous. It should be noted that the desirability of intervention in the purchase of liability insurance is different from the situation in the absence of the judgment-proof problem. As I emphasized in section 4, in the benchmark case without the judgment-proof problem, regulation of liability insurance coverage is not desirable.

Requirement to purchase coverage. Consider first a requirement to purchase complete liability insurance coverage. How will such a requirement affect incentives to reduce risk? If insurers can observe levels of care, a requirement to purchase coverage will lead to optimal incentives to reduce risk because premium reductions will reflect the full effect of care on expected harm. Suppose that an injurer with assets of only $100,000 who faces a $1,000,000 potential liability is required to buy complete liability insurance coverage and can reduce the risk of liability by 1 per cent by spending $5,000. He would receive a premium reduction of $10,000 for taking the step, so would be properly led to do so, but he might well not do so if he did not own coverage and had only $100,000 at risk, for then the expected value of the reduction would be only $1,000 to him.

However, if liability insurers cannot observe levels of care, a requirement to purchase complete coverage will tend to reduce incentives to take care. If the injurer with assets of $100,000 is required to purchase full coverage against the $1,000,000 risk, and insurers cannot observe his level of care, he will have no incentive at all to take care. Yet if he had not purchased coverage (or his coverage was less than $100,000), he would have a positive incentive to take

[132]The implications of the judgment-proof problem (and of escape from liability, see section 7.4 below) for the regulation of liability insurance were first addressed from an economic perspective in Shavell 1986, 1987a. See also Jost 1996, Polborn 1998, Shavell 2000, and Skogh 2000.

care, even though generally a suboptimal level of care. Thus, when insurers cannot observe levels of care, a requirement to purchase coverage tends to exacerbate the inadequacy of incentives due to the judgment-proof problem.

In any case, a requirement to purchase coverage has the beneficial aspect that it improves the activity level decision, because it confronts the injurer with the full social cost of his activity.[133] Otherwise, the activity level of injurers tends to be excessive, due to the judgment-proof problem.

A requirement to purchase coverage also has the advantage that it results in injurers being protected against risk. It should not, however, be considered to have the advantage that it protects victims against risk, because our assumption is that victims can purchase accident insurance coverage against risk.

The conclusion is that a requirement to purchase full liability insurance coverage is desirable if liability insurers can observe injurers' levels of care, for in that case the requirement results in optimal levels of care, and has other benefits as well, concerning activity levels and risk-bearing. But the requirement may not be desirable when insurers cannot observe levels of care, for in that case there will be a perverse effect on incentives to reduce risk.

Prohibition against coverage. Now let us examine the opposite policy, of prohibiting the purchase of liability insurance. This may increase the level of care that injurers would take if injurers would otherwise have purchased some liability coverage. For if their entire assets are exposed to risk, injurers may well take more care than otherwise. The resulting level of care might be an improvement, but it might also be excessive.

A prohibition against coverage will also tend to reduce activity levels from their too high levels in the absence of regulation. As with the level of care, the reduction in activity levels might be an improvement, but it might also be excessive.

Also, a prohibition against coverage has the disadvantage that it subjects injurers to greater risk than they would bear if they would have purchased any liability insurance coverage.

On balance, therefore, a prohibition against coverage might be beneficial (and superior to requiring coverage), but only if liability insurers are unable to observe levels of care. In that case, prohibiting coverage may increase levels of care and reduce activity levels. The potential social benefits of doing this might outweigh the disadvantages.

7.4 Comment on escape from liability and regulation of liability insurance. The arguments just discussed concerning regulation of liability insurance apply in part when injurers are not judgment-proof but nevertheless have inadequate incentives to take care because they escape liability with a probability (see the discussion in section 9 of chapter 4). If this is the case, then regulation of coverage may help to increase an otherwise too low effort to reduce risk. In particular, prohibiting coverage may do that and thus might be beneficial. (Requiring full coverage would in principle be a moot form of regulation, because parties would not rationally under-insure due to the chance of escaping liability; they would only do that due to the judgment-proof problem.)

---

[133]This is clearly the case under strict liability. Under the negligence rule, there is a beneficial activity level effect to the degree that injurers might be found negligent, as noted in section 2 of chapter 4.

7.5 Regulation of liability insurance in reality. Although, as stated in section 4.5, the sale of liability insurance is generally permitted, there are exceptions. Liability insurance is not allowed to be sold in some jurisdictions against punitive damages, against liability arising from certain willful acts, and against many criminal penalties.[134] The justification given is that this would tend to interfere with the public policy of deterring and punishing very undesirable behavior.[135] This justification comports with some of the points made in sections 7.3 and 7.4, in view of the possibility that deterrence may be inadequate because of the judgment-proof problem and the likelihood of escaping detection.[136] Liability insurance is also required in certain domains, notably for drivers of cars.[137] However, the chief justification given for such requirements is that they furnish implicit coverage to victims.[138] That seems a mistaken justification, as the rationale ought to be, as argued here, that the requirement would improve incentives to reduce risk, moderate activity levels, and provide insurance for injurers.

---

[134]See, for example, Jerry 1996, 471-77, Keeton 1971, 285-305, McNeely 1941, and Keeton, Dobbs, et al. 1984, 586.

[135] See, for example, Keeton, Dobbs, et al. 1984, 586, and Jerry 1996, 400, 472.

[136]This statement is speculative, in part because in some contexts it is not evident on a priori grounds why mandated liability coverage would not be a good way of increasing adherence to the law.

[137] See, for example, Keeton, Dobbs, et al. 1984, 601-03, and Jerry 1996, 859-63.

[138] See, for example, Keeton, Dobbs, et al. 1984, 600-03, and Jerry 1996, 860.

Chapter 6

Liability and Administrative Costs

In this chapter, I consider a third element, that of administrative costs, in evaluating the liability system. Administrative costs are the legal and other expenses borne by parties in resolving disputes that arise when harm occurs. I will first discuss the general nature and magnitude of administrative costs; as will be seen, they are of substantial importance. Then I will discuss the socially desirable use of the liability system given its administrative costs. Last, I will compare the private incentive to make use of the liability system to its socially desirable use.

1. Nature and Importance of Administrative Costs

1.1 Administrative costs described. As stated, administrative costs comprise the legal and other expenses that parties bear when accidents occur. These nonlegal expenses include the often significant time and effort of the involved parties as well as emotional costs and disutility. (There may also be utility enjoyed from the litigation process, especially by plaintiffs, and this must be set off against the costs of litigation.)

It is important to recognize that administrative costs are incurred not only with cases that go to trial, but also with cases that settle. Cases that settle involve substantial administrative costs because settlement may involve considerable time and negotiation, as well as the use of the legal system in motions and other actions that precede trial. Indeed, because more than 95 per cent of cases settle,[139] accounting for settled cases is of major significance, and it would be a mistake to attribute administrative costs mainly to the costs of trials.

What is the magnitude of administrative costs? Existing data suggest that in the United States the administrative costs of the liability system are large. Many studies find that administrative costs, averaged over settled and litigated claims, approach or exceed the amounts received by victims.[140] That is, for every dollar received by a victim, a dollar or more is spent

---

[139]Recent data on state courts show that in fiscal year 1992, over 96 percent of civil cases were settled or otherwise disposed without trial; see Ostrom, Kauder, and LaFountain 2001a, 29. Similarly, statistics on civil cases in U.S. district courts during the 2001 fiscal year show that almost 98 percent of cases were resolved without trial; see U. S. Department of Justice 2001a, 154. Cane 1999, 213, cites various studies suggesting that perhaps 99 percent of claims are settled without trial in the United Kingdom.

[140]A recent study, Tillinghast-Towers Perrin 2002, 12, reports that 42 per cent as victims' share of payments made by defendant parties in a nationwide survey of the tort system; in an earlier version of the study, Tillinghast-Towers Perrin 1995, 8, the figure was 46 per cent. Other studies include: Danzon 1985, 187, who reports 60 per cent as victims' share of medical malpractice liability insurance payments; Huber 1988, 151, who reports 40 per cent as victims' share of medical malpractice liability insurance payments, 50 per cent as victims' share of products liability insurance payments, and 50 per cent as their share of motor vehicle accident liability insurance payments; Kakalik et al. 1983, who estimate 37 per cent as victims' share of asbestos liability payments; Kakalik and Pace 1986, vii, who suggest from 45 per cent to 47 per cent as victims' share of tort payments; Keeton et al. 1983, 891, who review studies implying 44 per cent as victims' share of

delivering the dollar to him. It is not clear, however, to what extent these administrative costs should be viewed as intrinsic to the liability system or as a feature of the particular system that has developed.[141]

1.2 Margin over administrative costs of insurance. For some important purposes, it is not the total administrative costs of the liability system, but only their margin over the administrative costs of the accident insurance system that matters. If the liability system did not exist in some area of accident, those who would have received compensation under the liability system would often receive compensation through accident insurance. Therefore, the liability system increases administrative costs by the difference between its administrative costs and those of the insurance system.

The administrative costs associated with provision of accident insurance are much lower than those of the liability system, sometimes less than 10 percent of what victims receive.[142] The administrative costs of accident insurance are small by comparison because accident insurers have much less need than courts to inquire into the cause of losses or about injurers' behavior, because accident insurers have adopted comparatively simple procedures for verifying the magnitude of insureds' claims, and because accident insurers are not in an adversarial relationship with insureds.[143]

Because the accident insurance system is relatively inexpensive, the additional administrative costs of the liability system are quite high. Moreover, for victims who would not purchase accident insurance in the absence of liability (victims of low-magnitude harms, and large corporations), it is the total administrative cost of the liability system that should be considered to be its social cost, not its margin over insurance costs.

1.3 Strict liability versus negligence and administrative costs. The evaluation of strict liability and negligence rules depends on administrative costs as well as on incentives and risk-allocation. As will be seen, however, it is not clear on a priori grounds whether administrative costs will be higher under one rule than under the other, because of two conflicting considerations.

---

autombile accidents liability insurance payments. For the United Kingdom, Cane 1999, 397, reports 15 per cent as victims' share of tort compensation paid.

[141]I can readily imagine much less complex systems (for example, employing tabular schedules to decide damages). Also, the costs of the liability systems in other countries, such as France and Germany, may be significantly lower than in the United States.

[142]For example, the cost of administering the federal Old Age and Survivors Insurance Program is one half per cent of total expenditures and that of the Disability Insurance Program is 3 percent; see *Social Security Bulletin: Annual Statistical Supplement*, 2000, 18. The sales and administrative costs of commercial lines of all property and casualty coverage is about 25 per cent; see Fact Book 2001, 24; but most of this is probably sales costs, not administrative. For instance, for automobile coverage, administrative costs are 5 per cent and sales costs are 17 per cent; see Fact Book 2001, 50. And for homeowners policies, administrative costs are 6 per cent and sales costs are 22 per cent; see Fact Book 2001, 76.

[143]While there may, of course, be disagreement between insureds and insurers, it is unlikely to be as serious as that between victims and injurers. In part this is because insurers have an interest in honoring their policies: Private insurers will want to maintain their reputations; social insurers are presumably strongly motivated to serve insureds in any event.

First, the total number of claims is likely to be larger under strict liability than under the negligence rule, suggesting that administrative costs tend to be higher under strict liability. Under strict liability, a victim will have an incentive to make a claim whenever his losses exceed the costs of making a claim (assuming that he can credibly establish that the injurer was the cause of harm and that he himself was not contributorily negligent). Under the negligence rule, a victim will not have an incentive to make a claim so often because he will also be concerned about establishing the injurer's negligence. If a victim and an injurer both believe that a court will find the injurer free of fault, the victim will be unlikely to make a claim under the negligence rule.

But second, the average administrative cost per claim should be higher under the negligence rule. Under the negligence rule, it is more probable that a claim will be litigated than under strict liability, for under the negligence rule there is an additional element of dispute -- that of the injurer's negligence -- and hence more room for disagreement leading to trial. Because the probability of trial should be greater under the negligence rule and because trials will usually be more costly than settlements, we have one reason for saying that average administrative costs per claim are likely to be larger under the negligence rule. A second reason is that the costs of trial itself are likely to be higher under the negligence rule than under strict liability because the issue of negligence must be adjudicated under the former rule.

In sum, then, the comparison of the size of administrative costs under the two forms of liability is ambiguous as a theoretical matter. One would predict that a greater number of claims will be made under strict liability, but one would expect the average cost of resolving claims to be higher under the negligence rule because of both a higher propensity to go to trial and a higher cost per trial.

2. Socially Desirable Use of the Liability System Given Administrative Costs
2.1 The question whether the liability system is socially worthwhile.[144] Because of the administrative costs of the liability system, its use will be socially desirable if and only if its social benefits are sufficiently high. In particular, these social benefits concern the reduction in accidents (net of the costs of preventing them) but not compensation of risk-averse victims, given the assumption that accident insurance would be likely to furnish victims with compensation in the absence of the liability system. Thus, the use of the liability system will be socially worthwhile if and only if the savings from accident reduction it brings about exceed its administrative costs.[145]

A general implication is that the liability system will not be socially worthwhile if it would produce a sufficiently small expected reduction in accident losses. We would expect that to be true when the magnitude of accident losses is low, and thus for a whole range of rather

---

[144]The subject of the socially desirable amount of litigation, and of the possible divergence between this and the privately-determined amount of litigation (see the next section), will be considered in greater detail in section 2 of chapter 17.

[145]Of course, if the benefits of the liability system are properly conceived to be wider, this statement would be modified in a straightforward way.

trivial harms, such as bumping into someone when boarding a bus or insulting someone in a minor way. The reduction in accident losses following from liability would also be small where, even if the magnitude of possible harms is not low, there is little that individuals would do to prevent losses on account of liability. For example, it might be that liability for automobile accidents does not much affect the incidence of these accidents, for drivers' precautions may be determined largely by their fear of injury to themselves in accidents and by criminal liability for traffic offenses and for drunk driving.[146]

When, however, the liability system creates substantial incentives toward safety that do exceed administrative costs, the liability system will be socially worthwhile. Other things being equal, that would be so when the magnitude of possible losses is high, and when there are steps that will be taken to reduce them or their likelihood substantially.

2.2 Optimal payments by liable injurers, given administrative costs. Supposing that the liability system is worthwhile to employ, the presence of administrative costs affects the amount that liable injurers ought to pay because these costs raise the social costs of accidents: When an accident occurs that results in a legal dispute, the true social costs of the accident are comprised of its direct costs (harm to property and to person) *plus* administrative costs. The administrative costs are real social costs just as are the direct costs. For example, lawyers' services absorb valuable human resources, just as do physician services. Therefore, for injurers to have correct incentives to prevent harm, the amount that liable injurers ought to pay should equal the sum of the direct harm plus administrative costs. Now injurers naturally bear their own legal costs, time, and effort. Therefore, the optimal amount for injurers to pay equals the direct harm plus the legal and associated costs of victims and of the state. If the victim's harm is $100,000, his legal and other costs of handling the dispute are $30,000, and the court's costs are $5,000, the injurer's payment should be $135,000, not $100,000. For injurers' payments to be limited to $100,000 would lead to underdeterrence.

It should be noted that the point that injurers should pay for the direct harm plus the victim's and the state's costs does not mean that damages received by victims should equal that amount. The general argument made about victims' need for compensation and the optimality of fines (see section 6 of chapter 5) suggests that damages equal the direct (pecuniary) harm plus the victim's litigation costs, but that the state's costs be collected from the injurer in the form of a fine.

In fact, damages are not generally raised in the light of the victim's litigation costs or of the state's. Because the magnitude of victims' and the state's costs may average at least one third of direct harm,[147] damages appear to be inadequate for purposes of deterrence by a significant amount -- the situation is as if at least one third of conventional damages are ignored by the legal system.

---

[146]For a particularly direct and sustained expression by a torts scholar of the general view that the costs of the liability system may outweigh its incentive-related benefits, see Sugarman 1985.

[147]For example, in 1985 victims' legal payments as a fraction of harm was at least 33 per cent; see Hensler et al. 1987, 29 (in which harm is estimated by subtracting defendants' legal fees from their total payments).

3. Private versus Social Incentive to Use the Liability System Given Administrative Costs

      3.1 Private and social incentives may diverge. When a victim is injured, he will make a decision whether to sue, that is, whether to make use of the liability system, based on a comparison of his private benefits, and notably the expected judgment, and his legal costs. How are such private decisions concerning use of the liability system related to what is socially desirable, as discussed in section 2.1? The answer is that the private decision is quite different from the social one, and this leads to the possibility that the privately-generated use of the system could be either too large or too small relative to what is socially desirable.

      To illustrate the possibility that victims may bring a socially excessive number of claims, suppose that liability is strict, that there is absolutely nothing injurers can do to reduce risk, and that the social goal is to minimize the expected costs of accidents plus administrative costs. Clearly, in this example, it is socially undesirable for any claims at all to be made, for by assumption that cannot result in any reduction in accident losses, yet claims do result in administrative costs for victims and injurers. Nevertheless, claims obviously will be made: whenever a victim's loss exceeds his cost of making a claim, he will make a claim. The reason that claims are socially excessive here is evidently that victims' own financial return from making claims is often positive, even though the incentive thereby created is nil, a fact that is of no moment to a particular victim.

      Let me now illustrate the converse possibility that there may be too few claims, that is, that victims might not make claims even though, were they to do so, social welfare would be improved. Suppose that victims do not make claims because the cost of doing so, say 200, would exceed their loss of 100, which they frequently suffer. Knowing that they would not be sued, injurers have no incentive to reduce the high risk of these losses. But suppose that injurers could very cheaply take a precaution that would reduce the chance of the losses to almost zero. Then it would be socially desirable for victims to bring suit, because that would serve to induce injurers to reduce risk, and thereby also reduce the number of claims and administrative costs. The reason that the making of claims is socially inadequate in this example is that a victim's personal return from suit is negative; that the social value of suit is positive, due to its deterrent effects, is not of consequence to a victim contemplating suit.

      3.2 Divergence in incentives to bring suit may justify social intervention -- to limit or to subsidize use of the legal system. Because the volume of use of the legal system that results from private decisions to bring suit may diverge from what is socially best, social intervention will sometimes be necessary and useful.

      If the volume of litigation is too great because the administrative costs of suit outweigh the social benefits, reducing the use of the legal system will be worthwhile. For instance, it would be desirable to limit suits for automobile accidents if study showed that the deterrent effects of liability are modest (for the reasons noted above) in relation to the costs of the system. In such a case the state could contemplate such measures as imposing fees for bringing suit or banning suit altogether.

      Conversely, if the volume of litigation is inadequate, then subsidy or some other means of fostering suit may be desirable. For example, some types of low-magnitude harms that would not be privately worthwhile to pursue in view of the cost might be socially worthwhile to

promote due to the deterrent effect that the readiness of victims to sue would bring. To accomplish this, society could provide legal services for the bringing of the suits or otherwise support the process. Indeed, certain social efforts to promote access to the legal system, such as legal aid programs and small claims courts, might be rationalized in part along these lines.

The main point to stress is that the observed volume of litigation in any area of harm should not be viewed as approximately correct (in the way that the volume of some good sold in a normal market is so viewed). Because of the misalignment between social and private incentives to use the legal system, study of its benefits and costs is necessary to undertake to determine the direction and size of the divergence and the proper policy response to the divergence.

Note on the literature. In Shavell (1982b) I first examined the contrast discussed here between the socially desirable and the privately motivated use of the legal system in view of its costs.[148]

---

[148]See Menell 1983, Kaplow 1986b, Rose-Ackerman and Geistfeld 1987, and Shavell 1997, 1999, for further development of the social versus the private incentive to use the legal system..

References

ACLI Life Insurance Fact Book. 1999. Washington, D.C.: American Council of Life Insurance.

Arrow, Kenneth J. 1974. Optimal Insurance and Generalized Deductibles. *Scandinavian Actuarial Journal* 1974:1-42.

Beccaria, Cesare. [1767] 1995. *On Crimes and Punishments, and Other Writings*. Edited by Richard Bellamy, translated by Richard Davies, with Virginia Cox and Richard Bellamy. New York: Cambridge University Press.

Becker, Gary S. 1968. Crime and Punishment: An Economic Approach. *Journal of Political Economy* 76:169-217.

Bentham, Jeremy. [1789] 1973. *An Introduction to the Principles of Morals and Legislation*. In *The Utilitarians*. Reprint of 1823 edition. Garden City, N.Y.: Anchor Books.

Berman, Harold J. 1983. *Law and Revolution: The Formation of the Western Legal Tradition*. Cambridge, Mass: Harvard University Press.

Biggar, Darryl. 1995. A Model of Punitive Damages in Tort. *International Review of Law and Economics* 15:1-24.

Brown, John P. 1973. Toward an Economic Theory of Liability. *Journal of Legal Studies* 2:323-49.

Calabresi, Guido. 1970. *The Costs of Accidents: A Legal and Economic Analysis*. New Haven: Yale University Press.

----- 1975. Optimal Deterrence and Accidents. *Yale Law Journal* 84:656-71.

Calabresi, Guido, and A. Douglas Melamed. 1972. Property Rules, Liability Rules, and Inalienability: One View of the Cathedral. *Harvard Law Review* 85:1089-1128.

Calfee, John, and Richard Craswell. 1984. Some Effects of Uncertainty on Compliance with Legal Standards. *Virginia Law Review* 70:965-1003.

Cane, Peter. 1999. *Atiyah's Accidents, Compensation, and the Law*. Sixth edition. London: Butterworths.

Coase, Ronald H. 1960. The Problem of Social Cost. *Journal of Law and Economics* 3:1-44.

Cook, Philip J., and Daniel A. Graham. 1977. The Demand for Insurance and Protection: The Case of Irreplaceable Commodities. *Quarterly Journal of Economics* 91:143-56.

Cooper, Russel, and Thomas W. Ross. 1985. Product Warranties and Double Moral Hazard. *Rand Journal of Economics* 16:103-13.

Cooter, Robert D. 1982. Economic Analysis of Punitive Damages. *Southern California Law Review* 56:79-101.

----- 1989. Punitive Damages for Deterrence: When and How Much? *Alabama Law Review* 40:1143-96.

Craswell, Richard, and John E. Calfee. 1986. Deterrence and Uncertain Legal Standards. *Journal of Law, Economics, & Organization* 2:279-303.

Croley, Steven P., and Jon D. Hanson. 1993. Rescuing the Revolution: The Revived Case for Enterprise Liability. *Michigan Law Review* 91:683-797.

Cummins, J. David, Richard D. Phillips, and Mary A. Weiss. 2001. The Incentive Effects of No-Fault Automobile Insurance. *Journal of Law and Economics* 44:427-64.

Danzon, Patricia M. 1984. Tort Reform and the Role of Government in Private Insurance Markets. *Journal of Legal Studies* 13:517-49.

----- 1985. *Medical Malpractice: Theory, Evidence, and Public Policy*. Cambridge, Mass.: Harvard University Press.

Devlin, Rose Ann. 1990. Some Implications of No-fault Automobile Insurance. *International Review of Law and Economics* 10:193-205.

Dewees, Don, David Duff, and Michael Trebilcock. 1996. *Exploring the Domain of Accident Law: Taking the Facts Seriously.* New York: Oxford University Press.

Diamond, Peter A. 1974a. Accident Law and Resource Allocation. *Bell Journal of Economics* 5:366-405.

-----1974b. Single Activity Accidents. *Journal of Legal Studies* 3:107-64.

Dobbs, Dan B. 2000. *The Law of Torts*. St. Paul, Minn: West Group.

Donohue, John D, III. 1989. The Law and Economics of Tort Law: The Profound Revolution. (Review of *The Economic Structure of Tort Law,* by William M. Landes and Richard A. Posner, and *Economic Analysis of Accident Law,* by Steven Shavell.) *Harvard Law Review* 102:1047-73.

Eisenberg, Theodore, John Goerdt, Brian Ostrom, David Rottman, and Martin T. Wells. 1997. The Predictability of Punitive Damages. *Journal of Legal Studies* 26:623-61.

Eisner, Robert, and Robert H. Strotz. 1961. Flight Insurance and the Theory of Choice. *Journal of Political Economy* 69:355-68.

Emons, Winand. 1990. Efficient Liability Rules for an Economy with Non-identical Individuals. *Journal of Public Economics* 42:89-104.

Emons, Winand, and Joel Sobel. 1991. On the Effectiveness of Liability Rules when Agents Are Not Identical. *Review of Economic Studies* 58:375-90.

Eörsi, Gyula. 1983. Private and Governmental Liability for the Torts of Employees and Organs. In *International Encyclopedia of Comparative Law*, edited by André Tunc, vol. 11, ch. 4. Tübingen: J.C.B. Mohr.

Epple, Dennis, and Artur Raviv. 1978. Product Safety: Liability Rules, Market Structure, and Imperfect Information. *American Economic Review* 68:80-95.

Epstein, Richard A. 1980. *Modern Products Liability Law: A Legal Revolution*. Westport, Conn.: Quorum Books.

Fact Book 2001. 2001. New York: Insurance Information Institute.

Faure, Michael G. 2000. Environmental Regulation. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 2:433-520. Cheltenham: Edward Elgar.

Fleming, John G. 1998. *The Law of Torts*. Ninth edition. Syndey: LBC Information Services.

Geistfeld, Mark. 2000. Products Liability. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 3:347-95. Cheltenham: Edward Elgar.

Goldberg, Victor P. 1974. The Economics of Product Safety and Imperfect Information. *Bell Journal of Economics and Management Science* 5:683-88.

Grady, Mark F. 1983. A New Positive Economic Theory of Negligence. *Yale Law Journal* 92:799-829.

Green, Jerry. 1976. On the Optimal Structure of Liability Laws. *Bell Journal of Economics.* 7:553-574.

Grossman, Sanford J. 1981. The Informational Role of Warranties and Private Disclosure about Product Quality. *Journal of Law and Economics* 24:461-83.

Haddock, David D., Fred S. McChesney, and Menahem Spiegel. 1990. An Ordinary Economic Rationale for Extraordinary Legal Sanctions. *California Law Review* 78:1-51.

Halpern, Paul. 1998. Limited and Extended Liability Regimes. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 2:581-91. London: Macmillan.

Hamada, Koichi. 1976. Liability Rules and Income Distribution in Product Liability. *American Economic Review* 66:228-34.

Hansmann, Henry, and Reinier Kraakman. 1991. Toward Unlimited Shareholder Liability for Corporate Torts. *Yale Law Journal* 100:1879-1934.

Harsanyi, John C. 1977. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

Hart, H. L. A., and Tony Honoré. 1985. *Causation in the Law*. Second edition. Oxford: Clarendon Press.

Hensler, Deborah R., Mary E. Vaiana, James S. Kakalik, and Mark A. Peterson. 1987. *Trends in Tort Litigation*. R-3583-ICJ. Santa Monica: RAND Institute for Civil Justice.

Higgins, Richard S. 1978. Producers' Liability and Product-Related Accidents. *Journal of Legal Studies* 7:299-321.

Honoré, Anthony. 1983. Causation and Remoteness of Damage. Chapter 7 of volume 11, *Torts,* in *International Encyclopedia of Comparative Law*. Tübingen: J. C. B. Mohr.

Huber, Peter W. 1988. *Liability: The Legal Revolution and Its Consequences*. New York: Basic Books.

Huberman, Gur, David Mayers, and Clifford W. Smith, Jr. 1983. Optimal Insurance Policy Indemnity Schedules. *Bell Journal of Economics* 14:415-26.

Jerry, Robert H. 1996. *Understanding Insurance Law.* Second edition. New York: Matthew Bender.

Jolls, Christine, Cass R. Sunstein, and Richard Thaler. 1998. A Behavioral Approach to Law and Economics. *Stanford Law Review* 50:1471-1550.

Jost, Peter-J. 1996. Limited Liability and the Requirement to Purchase Insurance. *International Review of Law and Economics* 16:259-76.

Kahan, Marcel. 1989. Causation and Incentives to Take Care under the Negligence Rule. *Journal of Legal Studies* 18:427-47.

Kahneman, Daniel, Paul Slovic and Amos Tversky, editors. 1982. *Judgment Under Uncertainty: Heuristics and Biases.* New York: Cambridge University Press.

Kakalik, James S., and Nicholas M. Pace. 1986. *Costs and Compensation Paid in Tort Litigation.* Report R-3391-ICJ. Santa Monica: RAND Institute for Civil Justice.

Kakalik, James, Patricia Ebener, William Felstiner, and Michael Shanley. 1983. *Costs of Asbestos Litigation.* Report R-3042-ICJ. Santa Monica: RAND Institute for Civil Justice.

Kaplow, Louis. 1986b. Private versus Social Costs in Bringing Suit. *Journal of Legal Studies* 15:371-85.

Kaplow, Louis, and Steven Shavell. 1996b. Property Rules versus Liability Rules: An Economic Analysis. *Harvard Law Review* 109:713-90.

----- 2002c. On the Superiority of Corrective Taxes to Quantity Regulation. *American Law and Economics Review* 4:1-17.

Karpoff, Jonathan M., and John R. Lott, Jr. 1999. On the Determinants and Importance of Punitive Damage Awards. *Journal of Law and Economics* 42:527-73.

Keeton, Robert E. 1963. *Legal Cause in the Law of Torts*. Columbus: Ohio State University Press.

----- 1971. *Basic Text on Insurance Law*. St. Paul, Minn.: West Publishing Co.

Keeton, W. Page, Dan Dobbs, Robert Keeton, and David Owen. 1984. *Prosser and Keeton on the Law of Torts*. Fifth edition. St. Paul, Minn.: West Publishing Co.

Keeton, W. Page, Robert Keeton, Lewis Sargentich, and Henry Steiner. 1983. *Cases and Materials on Tort and Accident Law*. St. Paul, Minn.: West Publishing Co.

Keeton, William R., and Evan Kwerel. 1984. Externalities in Automobile Insurance and the Underinsured Driver Problem. *Journal of Law and Economics* 27:149-79.

Kessler, Daniel, and Mark McClellan. 1996. Do Doctors Practice Defensive Medicine? *Quarterly Journal of Economics* 111:353-90.

Kornhauser, Lewis A. 1982. An Economic Analysis of the Choice between Enterprise and Personal Liability for Accidents. *California Law Review* 70:1345-92.

Kornhauser, Lewis A., and Richard L. Revesz. 1998. Regulation of Hazardous Wastes. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 3:238-42. London: Macmillan.

Kraakman, Reinier H. 1986. Gatekeepers: The Anatomy of a Third-Party Enforcement Strategy. *Journal of Law, Economics, & Organization* 2:53-104.

Landes, Elisabeth M. 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, William M., and Richard A. Posner. 1981a. An Economic Theory of Intentional Torts. *International Review of Law and Economics* 1:127-54.

----- 1981b. The Positive Economic Theory of Tort Law. *Georgia Law Review* 15:851-924.

----- 1983. Causation in Tort Law: An Economic Approach. *Journal of Legal Studies* 12:109-34.

Larson, Arthur. 1994. *Larson's Worker's Compensation Law*. New York: Matthew Bender.

Le Gall, Jean-Pierre. 1983. Liability for Persons under Supervision. Chapter 3 of volume 11,

Torts, in *International Encyclopedia of Comparative Law*. Tübingen: J. C. B. Mohr.

Limpens, Jean, Robert Kruithof, and Anne Meinertzhagen-Limpens. 1983. Liability for One's Own Act. Chapter 2 of volume 11, *Torts,* in *International Encyclopedia of Comparative Law*. Tübingen: J. C. B. Mohr.

Litan, Robert E. 1991. The Safety and Innovation Effects of U.S. Liability Law: The Evidence. *American Economic Review: Papers and Proceedings* 81:59-64.

Markesinis, B. S. 1994. *A Comparative Introduction to the German Law of Tort*. Third edition. Oxford: Clarendon Press.

McCormick, Charles T. 1935. *Handbook on the Law of Damages*. St. Paul, Minn.: West Publishing Co.

McGregor, Harvey. 1983. Personal Injury and Death. Chapter 9 of volume 11, *Torts*, in *International Encyclopedia of Comparative Law*. Tübingen: J. C. B. Mohr.

McNeely, Mary C. 1941. Illegality as a Factor in Liability Insurance. *Columbia Law Review* 41:26-60.

Menell, Peter S. 1983. A Note on the Private versus Social Incentives to Sue in a Costly Legal System. *Journal of Legal Studies* 12:41-52.

----- 1998. Regulation of Toxic Substances. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 3:255-63. London: Macmillan.

Mill, John Stuart. [1861] 1998. *Utilitarianism*. Edited by Roger Crisp. Oxford: Oxford University Press.

Moore, Michael J., and W. Kip Viscusi. 1990. *Compensation Mechanisms for Job Risks: Wages, Workers' Compensation, and Product Liability*. Princeton: Princeton University Press.

O'Connell, Jeffrey, Phillip A. Bock, and Stewart Petoe. 1994. Blending Reform of Tort Liability and Health Insurance: A Necessary Mix. *Cornell Law Review* 79:1303-38.

Oi, Walter Y. 1973. The Economics of Product Safety. *Bell Journal of Economics and Management Science* 4:3-28.

Opoku, Kwame. 1972. Delictual Liability in German Law. *International and Comparative Law Quarterly* 21:230-69.

Ostrom, Brian J., Neil B. Kauder, and Robert C. LaFountain. 2001a. *Examining the Work of State Courts, 1999-2000*. Williamsburg, Va: National Center for State Courts.

Phillips, Jerry J. 1998. *Products Liability in a Nutshell*. Fifth Edition. St. Paul, Minn: West Group.

Pitchford, Rohan. 1995. How Liable Should a Lender Be? The Case of Judgment-Proof Firms and Environmental Risk. *American Economic Review* 85:1171-86.

-----. 1998. Judgment-proofness. In *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman, 2:380-83. London: Macmillan.

Polborn, Mattias K. 1998. Mandatory Insurance and the Judgment-Proof Problem. *International Review of Law and Economics* 18:141-46.

Polinsky, A. Mitchell. 1980c. Strict Liability vs. Negligence in a Market Setting. *American Economic Review: Papers and Proceedings* 70:363-67.

Polinsky, A. Mitchell, and William P. Rogerson. 1983. Products Liability, Consumer Misperceptions, and Market Power. *Bell Journal of Economics* 14:581-89.

Polinsky, A. Mitchell, and Steven Shavell. 1998b. Punitive Damages: An Economic Analysis. *Harvard Law Review* 111:869-962.

Pollock, Frederick, and Frederic William Maitland. 1911. *The History of English Law Before the Time of Edward I.* Second edition. Cambridge: Cambridge University Press.

Posner, Richard A. 1972a. *Economic Analysis of Law*. Boston: Little, Brown.

----- 1972b. A Theory of Negligence. *Journal of Legal Studies* 1:29-96.

----- 1973. Strict Liability: A Comment. *Journal of Legal Studies* 2:205-21.

Priest, George L. 1981. A Theory of the Consumer Warranty. *Yale Law Journal* 90:1297-1352.

----- 1988. Products Liability Law and the Accident Rate. In *Liability: Perspectives and Policy*, edited by Robert Litan and Clifford Winston, 184-222. Washington, D.C.: Brookings Institution.

----- 1991. The Modern Expansion of Tort Liability: Its Sources, Its Effects, and Its Reform. *Journal of Economic Perspectives* 5(no. 3):31 -50.

Restatement of the Law (Second) Torts. 1965. St. Paul, Minn.: American Law Institute.

Ringleb, Al H., and Steven N. Wiggins. 1990. Liability and Large-Scale, Long-Term Hazards. *Journal of Political Economy* 98:574-95.

Rose-Ackerman, Susan, and Mark Geistfeld. 1987. The Divergence Between Social and Private Incentives to Sue: A Comment on Shavell, Menell, and Kaplow. *Journal of Legal Studies* 16:483-91.

Rosenberg, David. 1984. The Casual Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System. *Harvard Law Review* 97:849-929.

Rubin, Paul H. 1993. *Tort Reform by Contract*. Washington, D.C.: AEI Press.

Rudden, Bernard. 1966. *Soviet Insurance Law*. Law in Eastern Europe Monograph No. 12. Leiden: A. W. Sijthoff.

Schäfer, Hans-Bernd. 2000. Tort Law: General. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 2:569-96. Cheltenham: Edward Elgar.

Schwartz, Gary T. 1992. The Beginning and the Possible End of the Rise of Modern American Tort Law. *Georgia Law Review* 26:601-702.

Shapo, Marshall S. 1994. *The Law of Products Liability*. Third edition. Salem, N.H.: Butterworth Legal Publishers.

Shavell, Steven. 1980a. An Analysis of Causation and the Scope of Liability in the Law of Torts. *Journal of Legal Studies* 9:463-516.

----- 1980b. Damage Measures for Breach of Contract. *Bell Journal of Economics* 11:466-90.

----- 1980c. Strict Liability versus Negligence. *Journal of Legal Studies* 9:1-25.

----- 1982a. On Liability and Insurance. *Bell Journal of Economics* 13:120-32.

----- 1982b. The Social versus the Private Incentive to Bring Suit in a Costly Legal System. *Journal of Legal Studies* 11:333-39.

----- 1985b. Uncertainty Over Causation and the Determination of Civil Liability. *Journal of Law and Economics* 28:587-609.

----- 1986. The Judgment Proof Problem. *International Review of Law and Economics* 6:45-58.

----- 1987a. *Economic Analysis of Accident Law*. Cambridge, Mass.: Harvard University Press.

----- 1997. The Fundamental Divergence Between the Private and the Social Motive to Use the Legal System. *Journal of Legal Studies* 26:575-612.

----- 1999. The Level of Litigation: Private versus Social Optimality of Suit and of Settlement. *International Review of Law and Economics* 19:99-115.

-----. 2000. On the Social Function and the Regulation of Liability Insurance. *Geneva Papers on Risk and Insurance: Issues and Practice* 25:166-79.

Skogh, Göran. 2000. Mandatory Insurance: Transaction Costs Analysis of Insurance. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 2:521-37. Cheltenham: Edward Elgar.

Sloan, Frank A. 1998. Automobile Accidents, Insurance, and Tort Liability. In *The New*

Palgrave Dictionary of Economics and the Law, edited by Peter Newman, 1:140-44. London: Macmillan.

Sloan, Frank A., Bridget A. Reilly, and Christoph M. Schenzler. 1994. Tort Liability versus Other Approaches for Deterring Careless Driving *International Review of Law and Economics* 14:53-71.

Spence, Michael. 1977. Consumer Misperceptions, Product Failure, and Producer Liability. *Review of Economic Studies* 44:561-72.

Statistical Abstract of the United States 2001. 2001. Economics and Statistics Administration, U.S. Census Bureau. Washington, D.C.: U.S. Department of Commerce.

Stoll, Hans. 1983. Consequences of Liability: Remedies. Chapter 8 of volume 11, *Torts,* in *International Encyclopedia of Comparative Law*, edited by André Tunc. Tübingen: J. C. B. Mohr.

Stone, Ferdinand F. 1983. Liability for Damage Caused by Things. Chapter 5 of volume 11, *Torts,* in *International Encyclopedia of Comparative Law*, edited by André Tunc. Tübingen: J. C. B. Mohr.

Sugarman, Stephen D. 1985. Doing Away with Tort Law. *California Law Review* 73:555-664.

Sykes, Alan O. 1984. The Economics of Vicarious Liability. *Yale Law Journal* 93:1231-80.

Tillinghast-Towers Perrin. 1995. *Tort Cost Trends: An International Perspective.* Chicago: Tillinghast-Towers Perrin.

----- 2002. *U.S. Tort Costs 2000: Trends and Findings on the Costs of the U.S. Tort System.* February 2002. Chicago: Tillinghast-Towers Perrin.

Tribe, Laurence H. 1971. Trial by Mathematics: Precision and Ritual in the Legal Process. *Harvard Law Review* 84:1329-93.

Tunc, André. 1983. Introduction. Chapter 1 of volume 11, *Torts,* in *International Encyclopedia of Comparative Law*. Tübingen: J. C. B. Mohr

Tversky, Amos, and Daniel Kahneman. 1974. Judgment under Uncertainty: Heuristics and Biases. *Science* 185:1124-31.

U.S. Census Bureau. 2001. *Health Insurance Coverage: 2000*. Current Population Reports, P60-215, September 2001. Washington D.C.: U.S. Census Bureau.

U. S. Department of Justice. 2001a. *Judicial Business of the United States Courts, 2001 Report of the Director*. Administrative Office of the United States Courts, Statistics Division. Washington, D.C.: U.S. Department of Justice.

U.S. Department of Labor. 2001. Employee Benefits in Private Industry, 1999. Release December 19, 2001. USDL 01-473. www.bls.gov/ncs/ebs/sp/ebnr0006.txt

U.S. Social Security Administration. 2000. *Annual Statistical Supplement, 2000, to the Social Security Bulletin*. Washington, D.C.: Social Security Administration.

Viscusi, W. Kip. 1983. *Risk by Choice: Regulating Health and Safety in the Workplace*. Cambridge, Mass.: Harvard University Press.

----- 1991. *Reforming Products Liability*. Cambridge, Mass.: Harvard University Press.

----- 1992. *Fatal Tradeoffs: Public and Private Responsibilities for Risk*. New York: Oxford University Press.

----- 1998. *Rational Risk Policy*. New York: Oxford University Press.

Viscusi, W. Kip, and William N. Evans. 1990. Utility Functions that Depend on Health Status: Estimates and Economic Implications. *American Economic Review* 80:353-74.

Von Bar, Christian. 1998. *The Common European Law of Torts*. Oxford: Clarendon Press.

Von Mehren, Arthur T., and James R. Gordley. 1977. *The Civil Law System: An Introduction to the Comparative Study of Law*. Second edition. Boston: Little Brown.

Wehrt, Klaus. 2000. Warranties. In *Encyclopedia of Law and Economics*, edited by Boudewijn Bouckaert and Gerrit De Geest, 3:179-99. Cheltenham: Edward Elgar.

Wittman, Donald. 1981. Optimal Pricing of Sequential Inputs: Last Clear Chance, Mitigation of Damages, and Related Doctrines in the Law. *Journal of Legal Studies* 10:65-91.

Zeckhauser, Richard J. 1973. Coverage for Catastrophic Illness. *Public Policy* 21:149-72.

Zweigert, Konrad, and Hein Kötz. 1998. *An Introduction to Comparative Law*. Third edition. Oxford: Clarendon Press.