CHAPTER 2

An Economic Model of Tort Law

This chapter and the next develop the economic model of accident, or tort, law. The model is based on the proposition that the rules of tort law are designed to give parties engaged in risky activities an incentive to undertake all reasonable means of minimizing the costs arising from those risks. For this reason, the economic model of accidents is usually referred to as the *model of precaution*. The purpose of this chapter is to develop this model in a general way so as to derive a set of basic principles that apply broadly to different areas of accident law. The next chapter then applies these results to specific areas.

The total costs of accidents consist of three components: the damages suffered by victims (in dollar terms); the cost of precautions against accidents by injurers and victims; and the administrative costs of the tort system. In this chapter, we focus on the first two of these costs as reflected in the model of precaution, while referring to administrative costs only in qualitative terms. In Chapter 8 we undertake a detailed analysis of administrative costs.

Although the model of precaution is outwardly a model of accidents, we will see in subsequent chapters that its usefulness extends beyond tort law to the areas of contracts and property. As such, it will be a useful tool for identifying connections across traditional legal boundaries.

1 What Is a Tort?

As noted, tort law is that area of the common law concerned with accidental injuries. Examples include personal injuries, products liability, workplace accidents, medical malpractice, and environmental accidents. As this list sug-

gests, risk is a necessary by-product of many socially beneficial activities, including driving, use of vaccines, medical procedures, and so on. And although we cannot ordinarily eliminate the risk without cutting out the activity altogether, we should nevertheless take all cost-justified steps to minimize the resulting cost. That means that we should invest in risk reduction to the point where saving an additional dollar in accident losses can only be achieved by spending more than one dollar in precaution.

Society has many ways of controlling risks, including safety regulation, taxation, and even criminal penalties for risky activities (for example, fines for speeding). These are all examples of "public" controls imposed by the government. This chapter is concerned instead with a private remedy—the right of accident victims to sue injurers for damages under tort law. (We consider the use of the public remedies just mentioned in later chapters.)

1.1 The Social Function of Tort Law

The primary social functions of tort law are twofold: to compensate victims for their injuries and to deter "unreasonably" risky behavior. Although the economic approach to tort law is not unconcerned with the goal of compensation, its primary goal is optimal deterrence. To this end, tort rules are viewed, first and foremost, as providing monetary incentives for individuals engaged in risky activities to take all reasonable (cost-justified) steps to minimize overall accident costs.

1.2 Elements of a Tort Claim

Since tort law is a private remedy for accidental harms, enforcement is in the hands of victims. In order to recover damages, a victim (plaintiff) must file a lawsuit against the injurer (defendant). (The fact that suits are costly has implications to be discussed below.) In order to prevail in the suit, the plaintiff has the burden of proving that the defendant is *legally responsible* and therefore must pay compensation. This requires that the plaintiff establish the following: (1) she sustained some damages; and (2) the defendant was the "cause" of those damages. In some cases, proving these two things is sufficient for the plaintiff to recover for her losses; in others, she must also prove "fault" on the part of the injurer. We will assume that the plaintiff succeeds in demonstrating damages and turn immediately to the second element, causation. Later, we examine the issue of fault.

Proving causation in a legal sense requires the plaintiff to establish two things: first, that the defendant's action was "cause-in-fact" of the damages; and second, that it was also "proximate cause." We examine these two notions of causation in turn.

1.2.1 Cause-in-Fact

Cause-in-fact is established by using the "but-for" test. Specifically, the plaintiff must prove that "but-for the defendant's action, the plaintiff would not have sustained any harm." In many cases, this is a straightforward matter: but for the explosion of the soda bottle, the plaintiff would not have been injured. In other cases, it is more problematic. One difficult case concerns two or more causes that simultaneously produce a harm that either would have caused acting separately. For example, suppose sparks from a train start a fire that combines with a fire set by a farmer clearing his land to burn another farmer's crops.[1] In that case, neither injurer's action will satisfy the but-for test because each can (correctly) claim that the damage would have occurred even if his fire had not gone out of control.

A second problematic situation arises when two or more injurers act to produce a harm that would *not* have occurred if each acted separately. For example, suppose that I push you backward while another person simultaneously pulls the chair out from under you, causing you to fall and injure yourself. In this case, *both* injurers satisfy the but-for test. The question then becomes how to apportion liability among the two injurers. For most of this chapter, we ignore these problems of causation by focusing on single-injurer accidents. (See Section 3.3.4 below and the discussion of multiple injurers in the context of environmental accidents in Chapter 3.)

A final problem with the but-for test is that it will often implicate extremely remote causes. As an illustration, consider the famous case of *Palsgraf v. Long Island RR.* (248 N.Y. 339, 162 N.E. 99, 1928). The facts of the case are as follows:

> Plaintiff was standing on a platform of defendant's railroad after buying a ticket to go to Rockaway Beach. A train stopped at the station, bound for another place. Two men ran forward to catch it. One of the men reached the platform of the car without mishap, though the train was already moving. The other man, carrying a package, jumped aboard the car, but seemed unsteady as if about to fall. A guard on the car, who held the door open, reached forward to help him in, and another guard on the platform pushed him from behind. In this act, the package was dislodged, and fell upon the rails. It was a package of small size, about fifteen inches long, and was covered with newspaper. In fact, it contained fireworks, but there was nothing in its appearance to give notice of its contents. The fireworks when they fell exploded. The shock of the explosion threw down some scales at the other end of the platform many feet away. The scales struck the plaintiff, causing injuries for which she sues.

Even though the railroad employee's action clearly satisfied the but-for test (that is, the accident would not have occurred but for the employee's actions),

the law did not hold the railroad liable. Limiting causation in this way is the purpose of proximate cause.

1.2.2 Proximate Cause

In addition to proving cause-in-fact, the plaintiff must also prove proximate, or legal, cause. That is, the connection between the injurer's action and the harm cannot be too remote. The usual test for proximate cause is the *reasonable foresight* test. In the *Palsgraf* case, for example, Judge Cardozo denied liability based on the argument that a reasonable person would not have foreseen that the railroad employee's action would result in harm to the victim.[2]

The foregoing shows the sorts of complications that arise from consideration of causation. Fortunately, we will see that the economic approach to tort liability simplifies the analysis of accidents to the extent that explicit considerations of causation can be largely ignored. Causation nevertheless plays such an integral role in actual tort cases that a positive economic theory of tort law cannot ignore it. We therefore return to this topic below.

1.3 Liability Rules

Once harm and causation have been established, the assignment of liability is determined by the application of a *liability rule*. A liability rule is simply a rule for dividing the damages between the injurer and the victim. Suppose, for example, that the victim has suffered damages of $10,000. A rule of *no liability* says that the victim should bear all of these costs herself. For example, in the area of products liability, the old rule of caveat emptor, or "buyer beware," is a form of no liability. In contrast, a rule of *strict liability* imposes all of the damages on the injurer; that is, the injurer must pay $10,000 to the victim. Strict liability therefore shifts liability from the victim to the injurer once causation is established.

A third type of liability rule, referred to as a *negligence rule*, shifts liability from the injurer to the victim only if the injurer is also found to be "at fault" or "negligent." A negligence rule is based on the idea that injurers owe potential victims a legal duty to take reasonable efforts to prevent accidents. If the injurer is judged by the court to have satisfied this duty, he is absolved from liability, even if he is the legal cause of the accident. However, if the injurer breached his duty, he is negligent and therefore liable for the victim's losses. In a sense we can think of negligence as a combination of no liability and strict liability, where the two are separated by a "threshold" based on the injurer's level of precaution. This way of thinking about negligence as a hybrid rule will prove useful in our discussion of other types of liability rules below.

2 An Economic Model of Accidents: The Model of Precaution

The analysis of tort law is based on the idea that legal rules for assigning liability are designed to minimize the total costs associated with risky activities. The basic model for doing this is referred to as the model of precaution.[3] In the simplest version (the unilateral care model), only the injurer can invest in costly precaution, or care, to reduce the likelihood and severity of the damages borne by the victim. In the more general model (bilateral care), both the injurer and victim can invest in precaution. In addition to the cost of precaution and damages, we need to take account of the administrative costs of using the legal system to resolve tort claims, including the costs of filing suit and conducting judicial proceedings. However, we initially ignore these costs and focus on precaution and damages.

We further simplify the analysis by assuming that injurers themselves suffer no damages. This could be relaxed, but doing so adds few additional insights.[4] Finally, we ignore for now the question of whether it is beneficial for the injurer and victim to be engaged in the risky activity and at what level. Below we explicitly consider this decision.

In analyzing the model, we first derive the socially efficient level of precaution, defined to be the level that would be chosen by a social planner. This outcome will then serve as a benchmark for examining the incentives created by actual legal rules.

2.1 The Unilateral Care Model

In developing the unilateral care model, we make use of the following notation:

x = dollar investment in precaution spent by the injurer;
$p(x)$ = probability of an accident;
$D(x)$ = dollar losses (damages) suffered by the victim.

We assume that both $p(x)$ and $D(x)$ are decreasing in x, reflecting the fact that greater precaution reduces both the probability and severity of an accident. Thus, expected damages, given by $p(x)D(x)$, are also decreasing in x. We further assume that they are decreasing at a decreasing rate. This means that precaution has a diminishing marginal benefit in terms of reducing accident risk. Intuitively, injurers invest first in the most effective precautions and only later turn to less effective measures.

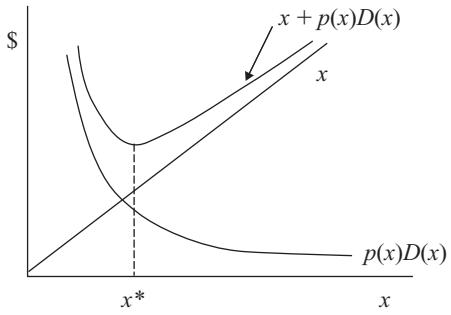

Figure 2.1
Social Costs in the Basic Accident Model

2.1.1 Social Optimum

The social problem, as noted above, is to choose x to minimize the costs of precaution plus expected damages. Formally, the problem is to

$$\text{minimize} \quad x + p(x)D(x). \tag{2.1}$$

The solution to this problem is best seen graphically in Figure 2.1, which graphs the cost of precaution (the positively sloped line), expected damages (the negatively sloped curve), and the summed costs (the U-shaped curve). The cost-minimizing level of care, labeled x^*, occurs at the minimum point of the total cost curve. At levels of care below x^*, an extra dollar of care reduces the victim's expected damages by more than one dollar, so total costs are reduced. However, beyond x^*, an extra dollar of care reduces expected damages by less than one dollar, so total costs rise.

Formally, x^* occurs at the point where the *slope* of the x curve equals the (negative) of the slope of the $p(x)D(x)$ curve. The slope of x reflects the *marginal cost of care* ($= \$1$), while the slope of $p(x)D(x)$ reflects the *marginal benefit of care* (the reduction in expected damages). (The optimal care level therefore does not necessarily occur at the intersection of these curves, which reflect total rather than marginal costs.)

2.1.2 Actual Care Choice by the Injurer

Consider now the actual choice of x by the injurer. To do this, we need to introduce the liability rule. First suppose the injurer faces no liability. In this case, the victim's damages are external to the injurer, so he simply minimizes his expenditure on precaution. Thus, he sets in $x = 0$, and total costs are not minimized.[5] (The injurer's private costs in Figure 2.1 therefore correspond to the x-line, which is minimized at the origin.)

Now suppose the rule is strict liability, such that the injurer is liable for the victim's full costs. (We ignore here problems in assigning a monetary value

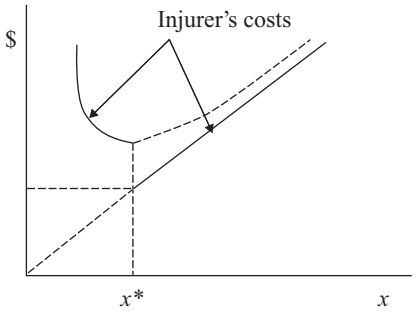

Figure 2.2
Injurer's Choice of Care Under a Negligence Rule

to a personal injury; see the discussion in Section 3.12 below.) In this case, the injurer will choose the socially optimal care level, x^*, because the threat of liability forces him to fully internalize the victim's expected damages. The injurer's costs now coincide with the social costs. Note that liability in this case functions exactly like a Pigovian tax in terms of its impact on injurer incentives (see the Appendix to Chapter 1), though it is different in that the victim rather than the government collects the revenue.

Finally, consider a negligence rule. Above we saw that under negligence law, the injurer owes a duty of reasonable care to all potential victims. If, in the event of an accident, the injurer is judged by the court to have met this duty, he avoids all liability (though he still must pay his cost of precaution). In contrast, if he breached this duty, he is fully liable. As noted above, negligence is therefore a combination of no liability and strict liability with the switch point at the due standard of care.

What is the due standard? The law talks about it in terms of the "reasonableness standard"—what level of care would a reasonable person undertake in the circumstances faced by the injurer? The economic theory of tort law equates the reasonableness standard with efficient care, or x^*. Below we discuss the justification for this equivalence. Here we focus on its impact on injurer incentives.

If x^* is the switch point between strict and no liability, then we can write the injurer's problem under negligence as follows

$$\text{minimize} \quad \begin{array}{ll} x, & \text{if } x \geq x^* \\ x + p(x)D(x), & \text{if } x < x^*. \end{array} \tag{2.2}$$

According to the first line of (2.2), the injurer can avoid liability for the victim's damages (and only pay his own cost of care) by just meeting the due standard. To see that this is the cost-minimizing strategy of the injurer, look at Figure 2.2, which is a version of Figure 2.1.

Note in particular that the injurer's costs coincide with social costs when he fails to take due care ($x < x^*$), but his costs are only the costs of precau-

tion when he meets (or exceeds) the due care standard ($x \geq x^*$). Clearly, the lowest point on the injurer's cost curve (the discontinuous solid line) occurs exactly at x^*.

Intuitively, if the injurer is taking less than due care, he benefits by increasing his care to x^* because by doing so, he *avoids all liability*. This is shown by the discrete drop in costs at the due standard. Further, the injurer gains nothing by raising x above x^*, but he must incur the additional costs of care. Thus, he will choose exactly x^*.

2.1.3 Comparison of Strict Liability and Negligence

The preceding shows that, in the unilateral care model, both strict liability and negligence result in efficient injurer care. However, this includes only the injurer's costs of precaution and the victim's damages. We noted above that we also care about administrative costs. Can we choose between the two rules on this basis?

First consider the cost per case. Strict liability will be cheaper to apply because plaintiffs need only prove causation, not fault. In contrast, in a negligence suit, the plaintiff will have to prove causation *and* fault. Thus, strict liability suits require less fact-finding and therefore involve less costly trials.

The cost per lawsuit is lower under strict liability, but there may be more suits. Remember that a tort claim must be initiated by the victim, who will only file a suit if the expected gain exceeds the cost. If a victim expects to lose, she will not file suit. Thus, under strict liability, the victim will file if (1) she can prove that the injurer caused her injuries, and (2) her losses exceed the cost of bringing suit.

Under negligence, the preceding conditions for filing must be met, but in addition, the victim must prove that the injurer is at fault (that is, that he failed to meet the due standard of care). And since we saw above that the injurer has a powerful incentive to meet the due standard, victims will often be deterred from filing suit under negligence. (Below we note some reasons why actual injurers may sometimes fail to meet the due standard.)

Thus, we expect fewer lawsuits under negligence as compared to strict liability. Taking this fact into account, we conclude that the calculation of overall litigation costs under the two rules is ambiguous: while strict liability likely leads to less costly suits, negligence leads to fewer overall suits. The comparison is therefore an empirical one.

Consider two other factors that, based on the analysis to this point, may affect the choice between strict liability and negligence. First, suppose there are errors in calculating the due standard of care (we examine this issue in more detail below). If the court systematically errs in setting the due stan-

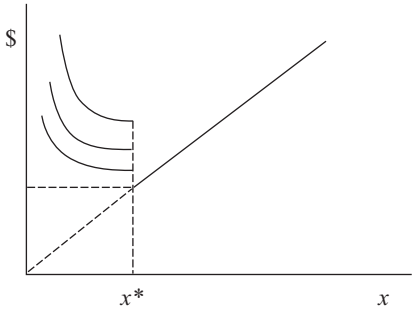

Figure 2.3
Injurer's Choice of Care When Victim's Damages Are Measured with Error

dard, it may result in an inefficient care choice by the injurer. This is not a problem under strict liabililty because the court need not calculate a due standard.

Suppose instead that the court makes errors in calculating the amount of the victim's damages. This will result in an inefficient care choice under strict liability, but so long as the error is not large, it will not distort the injurer's care choice under negligence (assuming that the due standard is set correctly). This is true because of the discontinuity in the injurer's costs under negligence, as shown in Figure 2.2. In particular, as Figure 2.3 shows, if damages are set too high or too low, the segment of the injurer's costs to the left of x^* shifts up or down, but so long as it doesn't shift down too much (that is, so long as the victim's damages are not underestimated by too much), the discontinuity remains and the injurer's cost-minimizing choice of care is x^*. A further basis for choosing between strict liability and negligence can be found by extending the accident model to allow for victim care.

2.2 Bilateral Care Model

We now make the above model more realistic by allowing victims as well as injurers to take care to reduce the likelihood and severity of an accident. For example, pedestrians decide which side of the street to walk on, and consumers of dangerous products decide whether to follow the manufacturer's safety instructions.

The above model only needs to be amended slightly to incorporate this change. Thus, we define:

$$y = \text{dollar investment in care by the victim;}$$
$$p(x, y) = \text{probability of an accident;}$$
$$D(x, y) = \text{damages suffered by the victim in the event of an accident.}$$

We now assume that expected damages, $p(x, y)D(x, y)$, are decreasing in both x and y.

The social problem in this case is to choose both x and y to

$$\text{minimize} \quad x + y + p(x, y)D(x, y). \tag{2.3}$$

Let x^* and y^* denote the resulting levels of care, both of which are assumed to be positive. (A graphical depiction of the optimum would require a three-dimensional analog to Figure 2.1.) Now consider the actual choices of x and y under the various liability rules.

2.2.1 No Liability and Strict Liability

We consider no liability and strict liability together because, in the context of the bilateral care model, they turn out to be mirror images of each other. (This symmetry demonstrates the sense in which "no liability" is in fact a liability rule.) Under no liability, the injurer bears none of the victim's damages and therefore, as in the unilateral care model, he invests in no care; that is $x = 0$. The victim, in contrast, fully bears her own damages and therefore chooses optimal care, y^*.[6] In the bilateral care model, no liability for the injurer could equivalently be called strict liability for the victim.

Correspondingly, under strict liability, the injurer faces full liability for the victim's damages and therefore chooses optimal care, x^*. In contrast, the victim is fully compensated for her losses (again, ignoring the problem of accurately measuring the victim's damages) and therefore bears none of her losses. She therefore chooses zero precaution, or $y = 0$. Thus, strict liability for the injurer is equivalent to no liability for the victim.

The preceding shows that in the bilateral care model, neither strict liability nor no liability lead to the efficient outcome. This illustrates a fundamental problem—namely, that both parties must face full responsibility for the damages at the margin in order to have the proper incentives. Otherwise, there is a *moral hazard problem* that results in too little precaution by one of the parties (or both if the damages are shared).[7]

Note that one way to achieve bilateral responsibility is to assess the injurer the full amount of the victim's damages, but then *not award the injurer's payment to the victim*. This will lead to the efficient outcome because the injurer will act as if the rule were strict liability, and the victim will act as if the rule were no liability. This is in fact the case under a Pigovian tax (and criminal fines), where the revenue from the tax is not used to compensate victims. The problem is that actual liability rules are not structured in this way; instead, they require that the victim receive whatever the injurer pays.[8] Of course, this reflects the compensatory function of tort law, but the analysis of the bilateral care problem suggests that this constraint (namely, that the victim must receive what the injurer pays) conflicts with its deterrence, or incentive function. It turns out that this is not true under negligence law.

2.2.2 Negligence

As discussed above, under negligence law, the injurer can avoid liability by meeting the due standard of care, x^*. This does not change when the victim also has the opportunity to take care. If the injurer chooses $x \geq x^*$, he avoids liability *regardless of the victim's choice of care*. Thus, as in the unilateral care model, the cost-minimizing strategy by the injurer is to meet the due standard.

Now consider the choice by the victim. Because she rationally anticipates that the injurer will meet the due standard, she expects to bear her own losses. Thus, she chooses her own care level, y, to

$$\text{minimize} \quad y + p(x^*, y)D(x^*, y). \tag{2.4}$$

Since she internalizes the full damages, she also chooses efficient care, y^*.

This shows that in a Nash equilibrium, a negligence rule with the due standard set at x^* induces both the injurer and victim to choose efficient care. Negligence therefore succeeds in achieving bilateral responsibility at the margin. The reason it can do this is that it employs two methods for inducing efficient behavior: first, it sets a threshold that allows the injurer to *avoid liability by meeting the threshold*; and second, it simultaneously imposes *actual liability* on the victim. We will encounter this use of a *threshold rule* for achieving bilateral responsibility (which is most clearly exemplified by the negligence rule) in other areas of the law.

Note the following aspects of this equilibrium. First, the victim is not compensated for her damages. This suggests that the compensatory and deterrence functions of tort law may be incompatible after all. (It turns out, however, that another threshold rule to be discussed below allows compensation of the victim while achieving bilateral efficiency.) Second, this equilibrium implies that no one is ever negligent—a result clearly at odds with reality. Below we discuss several reasons why parties may actually be negligent in equilibrium, including uncertainty over the due standard, differing costs of care across individuals, and limited injurer wealth.

2.3 The Hand Rule

Before extending the accident model in these and other directions, however, we first consider the extent to which actual negligence law corresponds to the economic ideal as just developed. In particular, we compare the legal definition of the due standard of care with the efficient standard, x^*. For purposes of this discussion, it is sufficient to restrict our attention to the unilateral care model.

The centerpiece of the positive economic theory of tort law—the argument that tort law embodies an economic logic—is the famous case of the

United States v. Carroll Towing Co. (159 F.2d 169, 2d Cir. 1947). The facts of the case are simple. A barge owner was accused of being negligent when he failed to post an attendant on board to make sure that the barge would not break loose from its moorings and cause damage to other ships and their cargo. The decision of the court was written by Judge Learned Hand, who wrote, in part:

> Since there are occasions when every vessel will break away from her moorings, and since, if she does, she becomes a menace to those about her; the owner's duty, as in other situations, to provide against resulting injuries is a function of three variables: (1) The probability that she will break away; (2) the gravity of the resulting injury, if she does; (3) the burden of adequate precautions. Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury, L; and the burden, B; liability depends upon whether B is less than L multiplied by P; i.e., whether $B < PL$.

The court ruled that in the circumstances at hand, the barge owner was in fact negligent for failing to post an attendant on board because the cost of doing so was less than the expected benefit, or $B < PL$. How does this simple inequality relate to the above model of accidents?

Recall that in the economic model, the due standard x^* was interpreted to be the level of injurer care that minimized the sum of the costs of precaution and expected damages. Thus, at x^*, the marginal cost of an additional unit of precaution equals the marginal benefit in terms of reduced damages. If we interpret B as the *marginal* cost of care and PL as the *marginal* reduction in accident costs from that last unit of care, then the injurer will be found negligent under the Hand rule if and only if $B < PL$, which is exactly the range over which $x < x^*$ in the economic model.

Now refer back to Figure 2.1 and recall that the curves reflected *total* rather than marginal costs and benefits of care. This is the reason that the optimal care level did not in general occur at the *intersection* of the curves. Rather, it occurred where the *slopes* of the curves were equal. Thus, the correct use of the Hand rule is based on the *slopes* of the x and pD curves. We will refer to this as the *marginal Hand rule*. One complication in applying marginal analysis to actual accident cases is that care usually does not vary continuously but comes in discrete bundles. The following exercise, based on the *Carroll Towing* case, illustrates the proper use of marginal analysis when applying the Hand rule to discrete care situations.

EXERCISE 2.1

Consider a barge owner who is deciding whether to post an attendant on his barge to make sure that it remains properly moored to the pier. The

following table gives the total cost of hiring the attendant, the probability of an accident, and the fixed cost of an accident:

	Cost of Care	*Probability*	*Damages*
No attendant posted	\$ 0	.25	\$400
Attendant posted for 24 hours	\$94	0	\$400

(a) Calculate the marginal cost, B, and marginal benefit, PL, of posting the attendant. According to the marginal Hand rule, would the barge owner be found negligent for failing to post an attendant?

Now suppose that the barge owner had a third option: post the attendant only during the day. The data for this option are as follows: Cost of care = \$50, Probability of an accident = .10, and Damages = \$400.

(b) Assume that the barge owner's only two options are "no attendant" and "post an attendant during the day." In this case, would the owner's failure to post an attendant be judged negligent by the marginal Hand rule?

(c) Assume that, prior to the accident, the owner had posted an attendant during the day. Suppose that the victim claims that the owner is negligent for not having posted the attendant for 24 hours. Use the marginal Hand rule to evaluate the merits of this claim.

(d) For each of the three options: "no attendant," "attendant during the day," and "attendant for 24 hours," calculate total expected costs (costs of care plus expected damages). Which option minimizes this total? Reconcile the result with your answers to (a)–(c).

2.4 The Reasonable-Person Standard

To this point we have treated all injurers as having identical costs of care (= \$1 per unit). In reality, injurers (and victims when they have an opportunity to take care) will have different costs of care, reflecting, for example, different ability levels, reflexes, or strengths. When costs of care differ, the cost-minimizing level of care will naturally be individual-specific. To see why, let c_j be the unit cost of care for individual j. Optimal care for that person will therefore minimize

$$c_j x + p(x)D(x). \tag{2.5}$$

Since c_j is the marginal cost of care in this case, individuals with higher values of c will have *lower* optimal care levels (given equal marginal benefits of

care across injurers). For example, if there are three types of injurers such that $c_1 < c_2 < c_3$, their cost-minimizing care levels will satisfy $x_1^* > x_2^* > x_3^*$. Intuitively, individuals with lower marginal costs of care should be held to higher due standards of care.

In general, however, the law does not individualize standards in this way. Rather, it sets a single standard applicable to all. This standard is based on a fictitious person referred to as the "reasonable person," who is defined to be "a personification of a community ideal of reasonable behavior, determined by the jury's social judgment. . . . Negligence is a failure to do what the reasonable person would do under the same or similar circumstances."[9]

The economic model implies that a single standard will not minimize overall accident costs when injurers differ in their marginal costs of care, so how can we explain the reasonable-person standard in the context of the economic model of tort law? The answer is that we have ignored administrative costs. Establishing what particular standard is appropriate in a given case would place a very high information burden on the court, a cost that is ordinarily too high compared to the savings in accident costs that would result from an individualized standard. (When the cost of individualizing the standard is low, however, the court will generally do so.)

What are the costs of setting a single standard, call it x^*, when injurers differ in their marginal costs of care? Two types of inefficiency result.[10]

1. For individuals with below-average marginal costs of care, the due standard is lower than what their individualized standard would be, or $x^* < x_j^*$. These individuals will have no incentive to increase their actual care above x^*, even though by doing so they would, by definition, lower social costs. (Remember that under a negligence rule, injurers avoid all liability by meeting the due standard, but they gain nothing privately by exceeding it.) All individuals with less than average costs of care therefore take too little care.
2. Individuals who have a cost of care that is slightly above average—and hence have an individualized standard less than the average (or $x^* > x_j^*$)—will actually increase their care level up to the due standard. This occurs because of the discontinuity in costs under the negligence rule, which creates a strong incentive for injurers to comply with the standard in order to avoid liability. Those injurers who find this privately beneficial, however, increase social costs by taking too much care (that is, the last dollar spent on care by these injurers reduces expected damages by less than one dollar).

There is a final group of injurers who actually choose the efficient level of care under the single standard. These are injurers whose cost of care is so high, and their individualized standard is so low relative to x^*, that they find it too costly

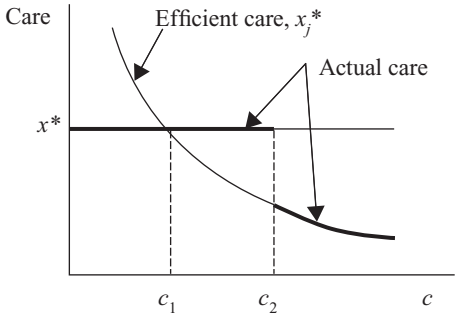

Figure 2.4
Comparison of Actual and Efficient Care Under a Single Due Care Standard

to raise their care level up to x^*. Instead, they choose x_j^* and are judged negligent. For these injurers, the negligence rule is equivalent to strict liability.[11]

Figure 2.4 shows the actual care choices of injurers (the darkened segments) compared to the level of care that minimizes accident costs (the downward-sloping curve). As the cost of care increases from left to right in the figure, we first see the set of injurers who take too little care ($c < c_1$), then those who take too much care ($c_1 < c < c_2$), and finally those who take efficient care and are found negligent ($c > c_2$). The fact that some injurers overinvest in care while others underinvest implies that we cannot say whether there will be more or fewer accidents compared to a rule with individualized standards (or under strict liability). However, total costs must be higher due to the inefficient care choices of the first two groups.

2.5 Contributory Negligence

Our discussion of the negligence rule has to this point focused on the legal duty of injurers to meet the standard of care. We have said nothing about a corresponding duty for victims, even though in our bilateral care model they can take care to avoid an accident as well. In fact, there is a form of negligence that is applied to victims; it is referred to as *contributory negligence*. Under a contributory negligence standard, victims are also required to meet a due standard of care as a condition for recovering for their injuries. For this reason, contributory negligence is a *defense* for injurers, which means that, even if an injurer admits to being negligent, he can still try to avoid liability by proving that the victim failed to meet the due standard (that is, was contributorily negligent).

Contributory negligence was first introduced in the old English case of *Butterfield v. Forrester* (11 East 60, K.B. 1809). The plaintiff in this case was injured while riding down a street when his horse collided with an obstruction that was negligently placed there by the defendant. The court held that, despite the defendant's negligence, the victim could not recover for his damages because of his own failure to act with due care. Specifically, the court said:

> A party is not to cast himself upon an obstruction which had been made by the fault of another, and avail himself of it, if he does not himself use common and ordinary caution to be in the right. . . . One person being in fault will not dispense with another's use of ordinary care for himself.

Contributory negligence can be paired with either a "simple" negligence rule or with strict liability. Let's examine it first when paired with simple negligence.

2.5.1 Negligence with Contributory Negligence

Under negligence with a defense of contributory negligence, the law establishes a due standard of care for both the injurer and the victim. Consistent with our analysis of simple negligence above, let these due standards be the efficient levels of care for the two parties—x^* for the injurer and y^* for the victim.

As noted above, contributory negligence bars recovery by the victim if she fails to meet her due standard of y^*, regardless of the injurer's choice of care. In contrast, if the victim chooses $y \geq y^*$, the injurer can still avoid liability by meeting his own due standard. Figure 2.5, panel (a), shows the assignment of liability for all choices of care by the injurer and victim under negligence with contributory negligence. Note that this differs from the assignment under simple negligence only in the lower left quadrant, the region where both parties are negligent. Under simple negligence (where the victim's choice of care is irrelevant), the injurer is liable in this case, while under negligence with contributory negligence, the victim is liable. In the other three quadrants, the rules are the same.

Does this change affect the efficiency of the negligence rule? The answer is no, provided that both due standards are set correctly. Consider first the choice of care by the injurer, and suppose that he expects the victim to satisfy the due care standard (the upper half of Figure 2.5[a]). In this case, the analysis is identical to that under simple negligence—the injurer chooses due

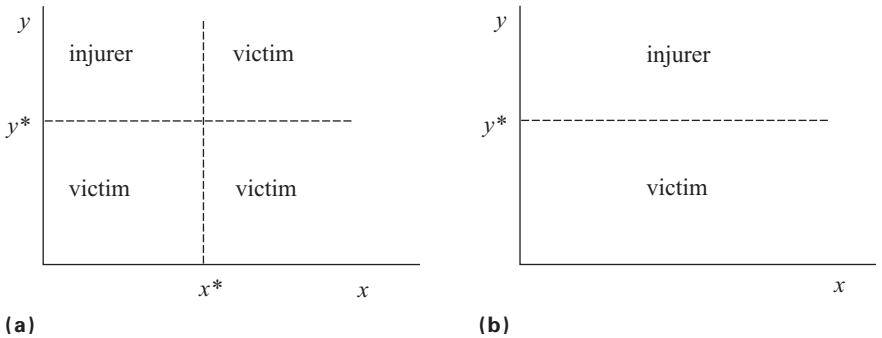

Figure 2.5
Assignment of Liability Under Negligence with Contributory Negligence (a) and Strict Liability with Contributory Negligence (b)

care to avoid liability. (Note that the answer is different if the injurer expects the victim to be negligent, for in that case, the injurer faces no liability and will therefore choose *zero care*. We consider this case in Section 3.1 below.)

Now consider the victim's incentives. If she expects the injurer to meet his due standard, the outcome is again identical to simple negligence—the victim bears her own losses and chooses efficient care of y^*. Thus, in a Nash equilibrium, both parties choose efficient care. This establishes that adding a defense of contributory negligence to simple negligence does not distort incentives. It also turns out not to affect the allocation of liability in equilibrium. Under both negligence rules, the victim bears liability in an efficient equilibrium.

According to the previous analysis, contributory negligence adds nothing to simple negligence in terms of either efficiency or allocation of liability. Further, it is likely a costlier rule to administer than simple negligence because it requires courts to evaluate compliance with two standards of care rather than one. Why then, until recently, has negligence with contributory negligence been the predominant tort rule in the United States? One possible reason will be illustrated below when we examine torts in which injurers and victims choose their care levels *sequentially* rather than simultaneously. However, we first examine contributory negligence when paired with strict liability.

2.5.2 Strict Liability with Contributory Negligence

Figure 2.5, panel (b), shows the assignment of liability under strict liability with contributory negligence. Note that in this case, only the victim's standard of care matters. In this respect, strict liability with contributory negligence is essentially a "negligence rule for victims." Thus, the Nash equilibrium is derived exactly as in the case of simple negligence, with the injurer and victim reversed. Specifically, the victim chooses due care to avoid liability, and the injurer, who is therefore strictly liable, chooses efficient care to minimize his costs.

Like the previous negligence rules, strict liability with contributory negligence therefore achieves bilateral efficiency, but it differs from the two negligence rules regarding the assignment of liability. Specifically, in an efficient equilibrium, the *injurer* bears the damages. This difference provides a basis for choosing between strict liability with contributory negligence and the two negligence rules on distributional grounds. If, for example, we determine, for policy reasons, that we want to favor victims as a group over injurers, but we don't want to distort incentives for efficient care, then we can employ strict liability with contributory negligence rather than negligence. (We will see a different reason for choosing strict liability with contributory negligence in our discussion of products liability in the next chapter.)

3 Further Topics

3.1 Sequential Care Accidents *

The bilateral care model to this point has been based on the assumption that injurers and victims make their care choices simultaneously, or, equivalently, that they make them without first observing the actual care choice of the other party. In this case, the parties had to form expectations about what the other party was doing. This actually helped promote efficiency because each party could act *as if* the other were taking due care. There is a substantial class of accidents, however, in which the injurer and victim move in sequence, and, as a result, the second mover can observe the *actual* care choice of the first mover before making his or her own choice. These are referred to as *sequential care accidents*.

In this type of setting, suppose the party moving first is observed to be negligent, due to inadvertence, error, or strategic behavior. Although the efficient outcome in which both parties take efficient care is now foreclosed, it is still desirable for the second party to take efficient care to avoid the accident. The question is whether the liability rules we have examined create such an incentive *once the second mover has observed the first mover's negligence*.[12]

3.1.1 The Injurer Moves First

Consider first the case where the injurer moves first. An example is provided by the facts of *Butterfield v. Forrester*. Recall that the injurer had negligently placed an obstruction in the street, and a passing rider collided with it and was injured. Assuming that the rider observed the obstruction in time to react, the question is whether the standard negligence rules provided him an incentive to take efficient steps to avoid the accident. Consider first simple negligence. According to the above analysis of the negligence rule, if the victim knows that the injurer has violated his due standard, then the victim has no incentive to take precautions because she knows that the injurer will be held liable for any damages. Thus, a simple negligence rule does not create incentives for victim precaution in the presence of observed injurer negligence.[13]

Now suppose a contributory negligence defense is added. Recall that under contributory negligence, negligent victims are liable regardless of the injurer's actions. Thus, even though the victim knows the injurer has been negligent, the victim nevertheless must take care in order to avoid liability. The conclusion is similar under strict liability with contributory negligence—once the victim takes due care, the injurer is strictly liable. This argument illustrates one advantage of adding a contributory negligence defense to simple negligence (Landes and Posner 1987, 76).

3.1.2 The Victim Moves First

We next consider the case of *Davies v. Mann* (10 M&W 546, 152 Eng. Rep. 588, 1842), which involves a similar situation to that in *Butterfield v. Forrester* except that the victim moves first. The victim in this case was the owner of a donkey that he had tied up next to a highway. The injurer subsequently drove this wagon down the highway and collided with the donkey, killing it. The court found that the owner of the donkey was negligent in having left it unattended on the side of the road, but the driver of the wagon was also found negligent because he was traveling at an excessive rate of speed.

Note that under a contributory negligence rule, the victim in this case would have been barred from recovery, in spite of the injurer's negligence. Thus, the injurer, who had time to observe the prior negligence of the victim, would have had no incentive to take reasonable care to avoid hitting the donkey. The above benefit of contributory negligence when the injurer moves first is therefore absent when the victim moves first. In contrast, a simple negligence rule would do better in this case because the injurer would have had to meet the due standard to avoid liability, regardless of the prior negligence of the victim.

3.1.3 Last Clear Chance

The preceding cases suggest that neither simple negligence nor negligence with contributory negligence can in all cases create incentives for the second mover in sequential accidents to take care in the presence of observed negligence by the first mover. The efficient rule depends on which party moves first. In response to this perceived deficiency, the court in *Davies v. Mann* articulated a rule that has since become known as *last clear chance*. Simply stated, the rule says that in sequential care accidents, the party acting second, whether the injurer or the victim, has the ultimate duty to exercise precaution against an accident, regardless of any prior negligence by the other party.

Note that in cases where the injurer moves first, like *Butterfield v. Forrester*, last clear chance is essentially equivalent to contributory negligence. If the injurer has acted negligently, both rules require the victim to take due care in order to avoid liability. However, in cases where the victim moves first, like *Davies v. Mann*, last clear chance is a necessary supplement to contributory negligence because it compels the injurer to take care despite the prior negligence of the victim. We therefore say that last clear chance "defeats" the injurer's attempt to use contributory negligence as a defense for his own negligence. As the court in *Davies v. Mann* recognized, were last clear chance not required of injurers, "a man might justify the driving over goods left on a pub-

lic highway, or even over a man lying asleep there, or purposely running against a carriage going on the wrong side of the road."

3.2 Comparative Negligence

All the liability rules that we have studied so far are what we call "all-or-nothing" rules. That is, one party, the injurer or the victim, bears all of the damages from an accident. In contrast, comparative negligence divides the damages between the injurer and victim in proportion to their relative fault. As of 1992, forty-four states had adopted some form of comparative negligence in place of standard negligence rules (Curran 1992). The principle reason for the conversion seems to have been a dissatisfaction with the perceived unfairness of all-or-nothing rules, especially in cases where, for example, slightly negligent victims are barred from recovering against grossly negligent injurers.

To illustrate the application of comparative negligence, consider a case in which a speeding motorist hits a pedestrian walking on the wrong side of the road. Suppose that the pedestrian incurs medical bills of $50,000. Under contributory negligence, she would be barred from recovering anything against the motorist, even if the court judged that 75 percent of the damage was due to his excessive speed. Under comparative negligence, in contrast, the injurer would be responsible for paying (.75) $\times$ ($50,000), or $37,500.

Although comparative negligence may be a fairer way of assigning liability for accidents, we need to ask whether this gain in fairness requires us to sacrifice the desirable efficiency properties of the all-or-nothing negligence rules. To answer this question, we examine the incentives for injurer and victim care under the most common form of comparative negligence, referred to as "pure" comparative negligence.[14]

As above, let x^* and y^* be the due standards of care for the injurer and victim, respectively. We can then define pure comparative negligence as follows: (1) if $x \geq x^*$ the injurer avoids all liability regardless of the victim's care choice; (2) if $x < x^*$ and $y \geq y^*$ the injurer is negligent and the victim is not, so the injurer bears full liability; and (3) if $x < x^*$ and $y < y^*$ both parties are negligent so they share liability in proportion to their fault. In the latter case, suppose that the injurer bears a fraction s of the damages, and the victim bears the remaining fraction $1 - s$, where s depends *positively* on the degree of injurer negligence and *negatively* on the degree of victim negligence.

Note that this rule and the two negligence rules (simple negligence and negligence with contributory negligence) differ from one another only in the assignment of liability when both parties are negligent (the lower-left quadrant in Figure 2.5[a]). The injurer bears liability in this case under simple negligence, the victim bears it under negligence with contributory negligence,

and the parties share it under comparative negligence. Thus, one can usefully think of simple negligence and negligence with contributory negligence as *special cases* of the more general comparative negligence rule. To see why, note that if we constrain $s = 1$, then the injurer bears full liability when both parties are negligent, as is true under simple negligence. In contrast, if we set $s = 0$, then the victim bears full liability when both parties are negligent, as under negligence with contributory negligence. Since both special cases provided efficient incentives for injurer and victim care, it is not surprising that the general rule can also be shown to provide efficient incentives. The proof is identical to that for the two negligence rules and is left as an exercise.

EXERCISE 2.2

Show that the comparative negligence rule as defined above results in an equilibrium in which both the injurer and victim take efficient care. To do this, first show that if the victim chooses due care of y^*, the best thing for the injurer to do is to choose x^*, and then show that if the injurer chooses x^*, the best thing for the victim to do is to choose y^*.

The fact that comparative negligence leads to an efficient equilibrium, and is fairer, suggests that it is superior to either of the all-or-nothing negligence rules. However, this is not necessarily true for two reasons. First, notice that in an efficient equilibrium, comparative negligence loses its desirable fairness properties because when both parties choose due care, the victim bears her own liability as under the other negligence rules. This is due to the threshold nature of the rule, which is the distinguishing feature of negligence rules, and the reason they are able to provide efficient bilateral incentives.

Second, comparative negligence has the drawback that it is probably costlier to administer than the other negligence rules because it requires the court to apportion damages based on relative fault. In many cases this will be a difficult task. Imagine, for example, trying to determine relative fault in a case where the customer at a drive-through restaurant spills hot coffee on herself while holding the cup between her legs. How much of the victim's damage was due to the coffee's having been too hot, and how much was due to the victim's mishandling of the cup?

Some economists have sought to demonstrate the superiority of comparative negligence over other forms of negligence by examining variations of the simple accident model. For example, they have shown that comparative negligence may be preferred when injurers and victims are risk averse (Landes and Posner 1987, 82), when there is uncertainty about the due standard of care (Cooter and Ulen 1986), when injurers differ in their costs of care (Ru-

binfeld 1987), or in the case of sequential care accidents (Rea 1987).[15]

Of course, only evidence from actual accidents can resolve the question of whether comparative negligence is more efficient than other negligence rules. White (1989) attempted to gather such evidence in the context of automobile accident cases in California from 1974 to 1976. (California switched from contributory negligence to comparative negligence in 1975.) Her results showed that contributory negligence created stronger incentives for accident avoidance, and further, that drivers took less than efficient care under comparative negligence. This suggests that the primary advantage of comparative negligence, in automobile accident cases at least, lies in its greater fairness.

3.3 Causation and Liability*

We return now to the issue of causation in relation to its impact on the assignment of liability. As noted above, issues of causation are often central to the actual assignment of liability in tort law, yet the economic model of accidents to this point has not explicitly raised the issue of causation. Cooter (1987a) has argued that this is because the economic model implicitly embodies a mathematical notion of causation through the functional relationship between precaution and expected damages. As a result, additional notions of causation are unnecessary to achieve efficient incentives for care. Nevertheless, a positive theory of tort law needs to address the court's use of causation principles in determining the scope of liability.

To keep the analysis simple, we focus on the unilateral care model and the simple negligence rule. Recall that in order to be held liable under negligence law, the injurer's failure to take due care must be both cause-in-fact and proximate cause of the victim's damages. We consider first the impact of the cause-in-fact requirement on the efficiency of the negligence rule.

3.3.1 Cause-in-Fact

To illustrate the impact of cause-in-fact on the negligence rule, consider the following example.[16] During a cricket game being played in a field enclosed by a 9-foot fence, a ball flies over the fence and injures a passerby. Suppose that the efficient height of the fence—the height that balances the cost of increasing its height against the savings in accident costs—is 10 feet. Based on the above characterization of the negligence rule, the owner's failure to build a 10-foot fence should therefore subject him to liability for *any* injuries suffered by a passerby.

Actual negligence law, however, does not operate in this way. According to the but-for test for causation, the owner would instead only be held liable for those accidents caused by his *negligence*; that is, for accidents caused by balls

TABLE **2.1** **Data for Cricket Example**

Height of fence (ft.)	*Cost of fence ($)*	*Accident costs ($)*	*Total costs ($)*
9	90	120	210
10	100	100	200
11	110	95	205

SOURCE: *Kahan (1989).*

TABLE **2.2** **Injurer's Costs Under Negligence Rule**

Height of fence (ft.)	*Cost under std. negligence rule ($)*	*Cost under negligence rule with cause-in-fact ($)*
9	210 = 90 + 120	110 = 90 + 20
10	100	100
11	110	110

SOURCE: *Adapted from Kahan (1989).*

that went over the 9-foot fence *but would not have gone over a 10-foot fence*. In other words, any balls that would have cleared a hypothetical 10-foot fence would not result in a claim for liability against the owner.

Does this restriction on liability eliminate the injurer's incentive to take due care under negligence? The answer is that it does not, though it does eliminate the discontinuity in the injurer's costs at the due standard of care (refer to Figure 2.2 above). The following numerical example, based on the cricket case, shows why. Table 2.1 shows the costs facing the owner of the cricket field, and Table 2.2 shows the owner's liabililty under a negligence rule, with and without the causation requirement, assuming that the due standard is a fence of 10 feet.[17]

Consider first the injurer's behavior under the standard negligence rule (column two in Table 2.2). If the owner builds a fence of less than 10 feet, he is negligent and therefore faces expected liability of $120, making his total expected costs $210. However, if he builds a fence of at least 10 feet, he is not negligent and hence faces only the cost of building the fence. His cost-minimizing choice is therefore to just meet the due standard of care by building the 10-foot fence at a cost of $100. In doing so, he expects to save the $120 in liability costs.

Now consider negligence with a cause-in-fact requirement (column three in Table 2.2). The only difference from the standard negligence rule is in the first row, where the injurer negligently builds a 9-foot fence. Although the injurer is liable for damages in this case, he is only liable for those damages caused by balls flying over the 9-foot fence that would not have cleared the 10-foot fence. Thus, his expected liability is the difference in expected accident

costs with a 9-foot fence compared to a 10-foot fence, or \$120 − \$100 = \$20, making his total expected costs from a 9-foot fence \$110. Note that this is still more than his cost from building a 10-foot fence, so the incentive for efficient care remains. The difference is that there is no longer a dramatic drop in the injurer's costs at the due standard of care.[18] This shows that, although the cause-in-fact requirement limits the injurer's liability under negligence, it does not distort his incentives to act efficiently.

It is important to emphasize that the preceding analysis does not provide an economic theory for the existence of the cause-in-fact limitation. It only showed that cause in fact is not inconsistent with efficiency. At the same time, however, it may eliminate the benefits associated with the discontinuity of the injurer's costs under a negligence rule.

3.3.2 Proximate Cause

In addition to proving that an injurer's negligence was cause-in-fact of an accident, the victim must prove that it was the proximate cause. Recall that the usual test for proximate cause is to ask whether the connection between the injurer's negligent act and the resulting accident was sufficiently close that a reasonable person, standing in the position of the injurer before the accident occurred, could have foreseen it. Proximate cause is therefore based on a "forward-looking" view of the accident, starting from the point in time when the injurer made his care choice. Note that this is in contrast to the backward-looking nature of cause-in-fact, which examines the causes of an accident with the benefit of hindsight.

The important consequence of the forward-looking nature of proximate cause is that the economic model of accidents takes the same perspective. Thus, we can use the apparatus of that model to construct an analytical version of the reasonable foresight test. To see how, consider an injurer whose actual care level was x', which is less than the due standard, x^*. Call $x^* - x'$ the "untaken precaution."[19] Assume that it has already been determined that the injurer's negligence was cause-in-fact of the accident. To determine if it was also proximate cause, the reasonable foresight test asks whether a reasonable person would have foreseen that his failure to meet the due standard would cause the victim's injuries. From the injurer's perspective before the accident has occurred, this amounts to asking how much his failure to exercise due care would *increase the expected damages to the victim*.

In terms of the economic model, the injurer's choice of x' rather than x^* increases expected damages by $p(x')D(x') - p(x^*)D(x^*)$. Note that this coincides with our definition above of the marginal benefit of increased care under the marginal Hand test, which we labeled PL. Under the reasonable foresight test, a finding of proximate cause requires that the increase in ex-

TABLE **2.3** **Data for Lifeguard Example**

Action	*Cost of care*	*Probability*	*Damages*	*Total costs*
No sign or lifeguard	0	.10	$1,000	$100
Sign	$5	.075	$1,000	$80
Lifeguard	$70	.005	$1,000	$75

pected damages due to the untaken precaution must exceed some threshold, call it T. That is, the injurer's negligence is proximate cause of the accident if $PL > T$, and it is not proximate cause if $PL < T$. Now, if we let $T = B$, the marginal cost of care, then *the test for proximate cause becomes identical to the marginal Hand test*.

It follows that the test for proximate cause and the test for negligence are in essence redundant tests. That is, both are forward-looking threshold tests for limiting the injurer's liability. On the one hand, this redundancy helps to explain why economic theories of negligence apparently have no need for causation principles (except for the notion of causation implicit in the functional relationship between care and expected damages). On the other hand, it again raises the question of why the law requires both inquiries before assigning liability. There are several possible reasons for including both tests.

First, proximate cause may serve to offset an inherent bias in the Hand test, which arises from the fact that in actual tort suits, the burden is on the plaintiff to propose the specific untaken precaution that constitutes negligence on the part of the injurer. To see the nature of this bias, consider an example based on the case of *Haft v. Lone Palm Hotel* (3 Cal.3d 756, 478 P.2d 465, 1970). The plaintiff in this case sought to recover damages when her husband and son, who were inexperienced swimmers, drowned in a hotel pool. Suppose that at the time of the drowning, there was no lifeguard on duty, nor was there a sign warning guests of this fact. Table 2.3 provides the data for this example.

As the table shows, a sign is inexpensive and reduces the probability of an accident slightly (from .10 to .075) by deterring some inexperienced swimmers, while posting a lifeguard is costly but reduces the probability significantly (from .10 to .005). The example assumes that posting a lifeguard minimizes total costs, so it is the efficient precaution. However, the plaintiff might find it easier to prove that failure to post a sign constituted negligence according to the marginal Hand test. Specifically, applying the test to the hotel's failure to post the sign, we find that $B = \$5$ and $PL = (.10 - .075)(\$1{,}000) = \$25$. Thus, the marginal savings in accident costs exceeds the marginal cost of care by a factor of five, so the test is easily satisfied. Now suppose that the

plaintiff proposed the failure to post a lifeguard as the untaken precaution. In this case, the marginal cost and benefit of posting a lifeguard are calculated relative to taking no action. Thus, $B = \$70$ and $PL = (.10 - .005)(\$1{,}000) = \$95$. The Hand test is therefore satisfied (as it must be since posting a lifeguard is efficient), but it is a much closer call as compared to the sign.

This example illustrates the incentive for plaintiffs to propose untaken precautions that are "too small" (that is, that fall short of the efficient precaution) in order to make it easier to satisfy the Hand test.[20] As a result, potential injurers looking at decisions in tort cases with this bias may perceive a due standard that is too low from a social perspective. A possible function of proximate cause in this setting, therefore, is to limit this *downward bias* in the Hand test by putting a lower bound on those untaken precautions that will pass the reasonable foresight test. In this case, for example, posting a sign reduces expected damages by $25, while posting a lifeguard reduces expected damages (relative to no action) by $95. Thus, by setting the threshold for the reasonable foresight test between these two values, the court can force the plaintiff to propose the efficient untaken precaution.

A second possible reason for requiring both the Hand test and causation may be to save on the administrative costs of using the legal system by limiting the scope of liability. That is, by requiring plaintiffs to prove both cause-in-fact and proximate cause, the law limits those circumstances in which a victim can recover against a negligent injurer (Shavell 1980a). This will deter some victims from filing suit, thereby saving on litigation costs. A third reason, also based on administrative costs, is that it may be intuitively easier for judges or juries to apply one or the other of the tests, depending on the facts of a particular case. That is, some cases may be easier to conceptualize in terms of cost-benefit principles, while others may be easier to see in terms of causation. Finally, it may simply be true that the causation requirement—both cause-in-fact and proximate cause—are designed to achieve goals other than efficiency such as fairness or distributive justice (Cooter 1987a).

3.3.3 *Res Ipsa Loquitur*

In some cases, the plaintiff may be unable to prove that the defendant's negligence was the cause of her harm, even though the circumstances of the accident make it exceedingly likely that it was. In these cases, the court may allow the plaintiff to recover, absent formal proof of causation, by invoking the doctrine of *res ipsa loquitur*, or "the thing speaks for itself." For example, in *Richenbacher v. California Packing Corporation* (250 Mass. 198, 145 N.E. 281, 1924), the plaintiff was allowed to recover against a food-packing company for damages suffered when she cut her mouth on glass found in a can

of spinach. Although there was no evidence of negligence in the defendant's packing operation, the court ruled that the only way the glass could have gotten into the can was by improper care during packaging.

When is it appropriate from an economic perspective for courts to apply the doctrine of *res ipsa loquitur*? Suppose that the accident technology is such that when the defendant employs due care, the probability of an accident is virtually zero (that is, $p(x^*) = 0$). In that case, the occurrence of an accident is necessarily evidence of negligence and also of cause-in-fact. In practice, however, this will not often be the case since efficient care usually does not prevent an accident with certainty. Note that if the doctrine is invoked in cases where $p(x^*) > 0$, then the liability rule effectively becomes strict liability because the injurer does not necessarily avoid liability by meeting the due standard. We will see in the next chapter that this was one route by which strict liability became the rule in products liability cases.

3.3.4 Uncertainty over Causation

Another circumstance in which a plaintiff might have difficulty proving causation is when there are multiple possible causes of her injuries (Shavell 1985). An example is when exposure to a toxic substance increases the "background" risk of developing cancer. If the victim develops cancer, it is not known whether it is due to the exposure or to a "natural" cause. Another source of uncertainty over causation is when there are multiple injurers. For example, suppose that two hunters both fire in the direction of a third party, but only one bullet strikes the victim.[21] In these cases, how should liability be assigned to the injurer(s) in order to induce efficient precaution?

To answer this question, we consider the case of a single injurer coupled with a background risk. (We consider the case of multiple injurers in the next chapter in the context of environmental accidents.) Specifically, let the probability of an accident be $p(x) + q$, where q is the background risk. For simplicity, let the victim's damages in the event of an accident be fixed at D. Since the background risk is constant, the social problem is to choose x to minimize expected costs, or to

$$\text{minimize} \quad x + [p(x) + q]D. \tag{2.6}$$

The resulting efficient level of precaution is x^*. In terms of the injurer's incentives, note that either a strict liability rule or a simple negligence rule with the due standard set at x^* continues to yield the correct incentives. To see this, rewrite expected costs in (2.6) as $x + p(x)D + qD$. Since the term qD is additive, it does not affect the minimum point. Thus, the background risk has no effect on incentives, and the problem becomes identical to the unilateral care model. In this respect, uncertainty over causation does not present a problem

regarding efficient care. However, it does potentially subject the injurer to liability for injuries that he did not "cause." In addition to being unfair, this increases administrative costs by expanding the scope of liability, and also may cause injurers to reduce their level of activity to an inefficiently low level (see the next section).

There are two ways to limit the scope of liability in the case of uncertainty over causation. The first is to employ a proximate cause limitation that holds the injurer liable only if the *conditional probability* that he caused the injuries exceeds some threshold. In the background risk model, this conditional probability is given by $p(x)/[p(x) + q]$. If the rule is strict liability, then the injurer's problem with the proximate cause limitation is

$$\text{minimize} \quad \begin{array}{ll} x, & p(x)/[p(x) + q] \leq T \\ x + [p(x) + q]D, & p(x)/[p(x) + q] > T \end{array} \tag{2.7}$$

for some threshold T. Now suppose that we set $T = p(x^*)/[p(x^*) + q]$. Since $p(x)/[p(x) + q]$ is decreasing in x (that is, the injurer's conditional probability of causation is decreasing in his care level), the condition for the injurer to avoid liability (the top line in [2.7]) is just $x \geq x^*$. The injurer's problem in (2.7) therefore becomes identical to that under simple negligence, and the injurer chooses efficient care.[22]

The second way that the injurer's liability can be limited in this case is to hold him strictly liable in all cases, but for less than the full amount of the victim's damages. Specifically, suppose that in the event of an injury to the victim, the fraction of the damages that the injurer must pay is equal to the conditional probability that he caused the damages. The injurer's problem in this case is to

$$\text{minimize} \quad x + [p(x) + q] \times \left(\frac{p(x)}{p(x) + q}\right)D. \tag{2.8}$$

Note that this is equivalent to minimizing $x + p(x)D$, which is has the solution x^*. Thus, this sort of proportional liability also induces the injurer to take efficient care.

A court actually employed a rule of this sort in a well-known case involving several manufacturers of a drug that was later found to be a cause of cancer.[23] Because of the time lapse between the purchase of the drug and the discovery of its carcinogenic nature, the plaintiff did not know which of several companies had sold the drug to her. The court therefore apportioned liability among the companies according to their market shares at the time of the sale, the latter serving as a proxy for the probability that each was responsible for the victim's damages.

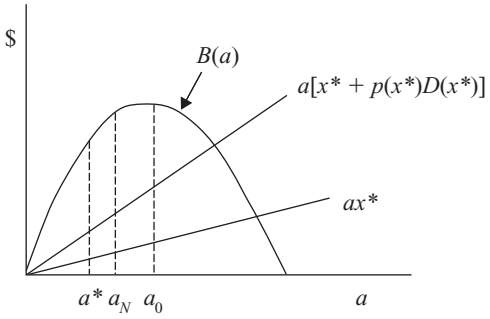

Figure 2.6
Injurer Activity Levels Under Different Liability Rules

3.4 Activity Levels

To this point, we have focused on injurer and victim *precaution* as the primary determinant of the probability and severity of accidents, but in many cases an equally important factor is their *activity levels*. That is, how intensively do the parties engage in the risky activity? For example, a motorist decides how fast to drive and how often to have his car inspected (both measures of care), but also how many miles to drive (his activity level). Similarly, the manufacturer of a dangerous product decides what safety features to include, as well as how many units to sell.[24] We illustrate the role of activity levels in the context of the unilateral care model and then generalize the results to the bilateral care model below.[25]

Let the injurer's activity level be denoted a, which yields him benefits of $B(a)$. Assume that $B(a)$ is a single-peaked curve that is maximized at a unique activity level a_0, as shown in Figure 2.6. Thus, in the absence of any accident risk, this is the level of activity that the injurer would choose, and it is also the efficient level.

We next need to specify the impact of the injurer's activity on accident costs. Assume that expected damages and costs of care are proportional to the injurer's level of activity. Thus, for example, if the motorist drives twice as many miles, his cost of precaution and the expected damages to victims both double.[26] Given this specification of the accident technology, we write total expected accident costs as $a[x + p(x)D(x)]$. The social problem is now to choose the injurer's level of activity and care to maximize net benefits, given by

$$B(a) - a[x + p(x)D(x)]. \tag{2.9}$$

Note that this problem can be broken into two parts. The first is to choose the level of care to minimize expected accident costs. Given the proportionality of accident costs to the activity level, it turns out that the optimal care level, x^*, is independent of a. Intuitively, the injurer should simply replicate his optimal care choice each time that he engages in the activity. For example, the

motorist should drive carefully on each trip, and the manufacturer should make each unit of output equally safe.

The second part of the problem is to choose the optimal activity level, a^*, given optimal care. This choice is shown graphically in Figure 2.6, where the optimal activity level occurs at the point of greatest vertical distance between the $B(a)$ curve and the ray representing expected accident costs, $a[x^* + p(x^*)D(x^*)]$. Equivalently, it occurs where the slopes of the two curves are equal, or where the marginal benefit of engaging in the activity equals the marginal accident costs. Note that the socially optimal activity level is less than the level that maximizes gross benefits $B(a)$ ($a^* < a_0$) because the latter does not take account of accident costs.

Now consider the injurer's choice of care and activity level under different liability rules. First note that under a rule of no liability, the injurer will choose an excessive activity level of a_0 (and also zero care) since he ignores the victim's damages. Next, consider the rules of strict liability and negligence. Recall that in the unilateral care model, both rules induce the injurer to take efficient care of x^*. Under strict liability the injurer will also choose the efficient activity level, a^*, since he fully internalizes the victim's damages. That is, his private benefits coincide with social benefits.

Under a negligence rule, however, he will not choose the efficient activity level. To see why, recall that once he meets the due standard of care, he avoids all liability for the victim's damages, though he does bear his costs of care. Thus, the injurer will choose his activity level to maximize $B(a) - ax^*$. As shown in Figure 2.6, this results in an activity level of a_N, which is too high from a social perspective (though it is not as high as under a rule of no liability). This shows that when activity levels matter, *strict liability is preferred to negligence in the unilateral care model*.

How does this conclusion extend to the bilateral care model in which victims as well as injurers can choose care and activity levels? To answer this question we first need to understand more clearly why the negligence rule induces efficient care but too much activity by the injurer. The reason is that it sets a due standard of care that allows the injurer to *avoid liability* by meeting the standard. While we have seen that this provides a powerful incentive for the injurer to comply with the standard with respect to care, it results in excessive activity precisely because the injurer does not internalize the full cost of his activity. In contrast, the injurer chooses efficient activity under strict liability because he does face full liability for the victim's damages. The general principle is that a party will choose the efficient activity level *only if he faces the residual damages from the accident*.[27] And since the liability rules we have studied impose actual damages on only one of the parties, it follows that none of them can simultaneously induce efficient activity levels by both parties.

Consider, for example, simple negligence, negligence with contributory negligence, and strict liability with contributory negligence. As we have seen, all three rules induce efficient precaution by both the injurer and the victim. Under the first two rules, victims will also choose the efficient activity level since they bear the residual liability, while injurers will choose an excessive level of activity. The reverse is true under strict liability with contributory negligence because under this rule, the injurer bears the residual liabiltity. Since none of the rules we have studied yields the efficient outcome along all four dimensions (the choice of care and activity by the injurer and victim), the best rule depends on a comparison of overall accident costs under each of the rules.

An important example of an injurer's activity level is the number of units of a dangerous product sold by the manufacturer (as distinct from the safety of each unit of the product). We will see in the next chapter that when the victims of an accident are customers of the injurer, some of the conclusions reached in this section regarding activity levels need to be altered.

3.5 Punitive Damages

To this point, our discussion has focused exclusively on *compensatory damages*, which are aimed solely at compensating victims' losses. However, in cases where the injurer's actions are seen as intentional or reckless, the court may also award *punitive damages*. As the name suggests, punitive damages are intended to punish the injurer for some perceived wrongdoing, as well as to deter future injurers from engaging in similar actions.[28] In this sense, punitive damages are similar to fines in criminal law (see Chapter 9).

The economic theory of punitive damages is based solely on the deterrence motive; that is, the desire to provide injurers with the correct incentives for care.[29] Our analysis of the accident model to this point, however, has shown that compensatory damages alone are sufficient to achieve this goal. It follows that adding punitive damages will actually result in *excessive* deterrence (too much care by injurers). What this conclusion ignores is that injurers may sometimes be able to escape liability for damages that they caused. One reason is the problem of uncertainty over causation discussed above—in some cases victims may have difficulty in identifying or proving the specific cause of their injuries. A second reason is that the cost of litigation may prevent some victims from bringing suit to collect damages.[30] Finally, injurers may sometimes take conscious steps to conceal their identity, especially when the injury was inflicted intentionally.

For these reasons, injurers may not expect to face the full damages that they cause and will therefore take too little care. Punitive damages address this problem by increasing the amount of damages injurers expect to pay in those cases where victims succeed in recovering damages. By appropriately

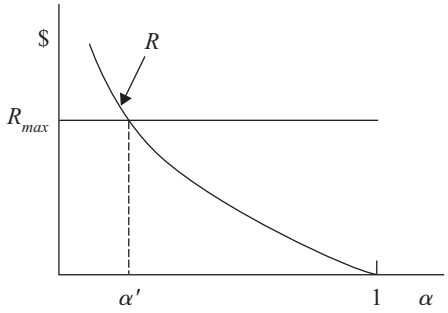

Figure 2.7
The Level of Punitive Damages that Achieves Efficient Deterrence

specifying the amount of these damages, courts can restore efficient incentives for injurer care.

We illustrate this in the context of the unilateral care model with a strict liability rule. Assume that an injurer expects to face liability for only a fraction α of the damages he causes, where $\alpha < 1$. In choosing his care, we will therefore minimize

$$x + p(x)\alpha D(x). \tag{2.10}$$

Since his expected liability is less than the full damages he causes, $p(x)D(x)$, the injurer will take less than efficient care. Further, the lower is α, the lower will be his care choice.

Now suppose that courts are able to award victims compensatory damages of $D(x)$ plus punitive damages of R, making the injurer's overall expected liability equal to $p(x)\alpha[D(x) + R]$. Incentives for efficient care are achieved when the injurer's expected liability equals the full expected damages of the victim, or when $p(x)\alpha[D(x) + R] = p(x)D(x)$. Solving this equation for R yields

$$R = \frac{1 - \alpha}{\alpha} D(x). \tag{2.11}$$

The efficient level of punitive damages is thus proportional to actual damages, where the factor of proportionality is given by $(1 - \alpha)/\alpha$. This factor is sometimes referred to as the punitive multiplier. It follows immediately from (2.11) that the amount of punitive damages is decreasing in α and equals zero when there is no risk of the injurer's escaping liability (that is, when $\alpha = 1$). Figure 2.7 graphs R as a function of α.

How closely do courts follow the above theory in calculating actual punitive damage awards? Based on their analysis of punitive damages, Mitchell Polinsky and Steven Shavell (1998, 898–99) conclude that they do not follow it very closely. In particular, they conclude, "Courts . . . do not pay systematic attention to the probability of escaping liability, even though this is

the central element in determining the appropriate damages multiplier for the purposes of achieving proper deterrence."

EXERCISE 2.3

Suppose that an injurer causes $500,000 in damages to a victim, but only faces a one-in-three chance of being found liable.

(a) Calculate the punitive multiplier.
(b) Calculate the amount of compensatory damages and the amount of punitive damages that a court should award if the victim brings suit. What is the injurer's overall liability?

Should Punitive Damages Be Capped? Excessive punitive damage awards in high profile tort cases often lead policymakers to propose caps on punitive damage awards. Indeed, many states have enacted such caps. The usual argument in favor of caps is that they limit incentives to file frivolous claims, thereby saving on administrative costs. Though this argument has some merit, there are two counterarguments, one theoretical and one empirical. The theoretical argument is that arbitrarily set caps on punitive damages may inhibit the deterrence function of punitive damages. For example, suppose a cap is set at R_{max} in Figure 2.7. This will have no effect on the ability of courts to achieve efficient deterrence when $R < R_{max}$, but in the range where $R > R_{max}$ (those cases where $\alpha < \alpha'$ in Figure 2.7), the cap will result in underdeterrence. Any benefits of a cap in terms of saved litigation costs must therefore be weighed against the cost of underdeterrence.

The empirical argument against caps is that punitive damages are not frequently awarded, and when they are, they are often overturned or reduced on appeal.[31] Thus, the popular perception of excessive awards, which is primarily based on a few high profile cases, apparently is not reflective of the overall population of cases.

3.6 The Judgment-Proof Problem

In some cases, defendants who are found liable have insufficient assets to pay the victim's damages. When an injurer has limited assets, we say that he is "judgment proof" (Shavell 1986). For example, a manufacturer of a dangerous product may go bankrupt before an accident occurs. The problem is that, if a potential injurer anticipates that he will be judgment proof in the future, he may take too little precaution in the present to avoid accidents. To illustrate, suppose that at the time he makes his care choice, an injurer expects to

be solvent in the future with probability α and insolvent (or judgment proof) with probability $1 - \alpha$. (Equivalently, the injurer expects to have assets equal to a fraction α of the victim's expected damages.) Note that, under a rule of strict liability, the injurer's problem in this case is identical to that in (2.10); as a result, he takes too little care.

The outcome may be different under a negligence rule. In particular, if the probability of being judgment proof is not too large (that is, if α is not too small), the injurer will still find it optimal to meet the due standard and avoid *all* liability.[32] The discontinuity in injurer costs under negligence thus helps to counteract the judgment-proof problem.

The fact that injurers may be able to avoid liability costs by being found judgment proof creates an incentive for firms to act strategically by, for example, divesting themselves of risky activities and locating them in small subsidiary firms, given the limited liability of assets within a corporation (Ringleb and Wiggins 1990). This may be privately profitable for the firm, but it distorts incentives for care as well as for the organizational structure of firms. As a result, if there is even a slight chance that the parent company will be held "vicariously liable" for the subsidiary's negligence, the expected costs of subcontracting may outweigh the benefits. In fact, Brooks (2002) found that oil companies actually decreased their use of independent shippers following the Exxon *Valdez* oil spill because subsequent legislation greatly increased the risk of vicarious liability.

The preceding analysis of the judgment-proof problem assumed that the injurer's asset level limited the amount he could pay in liability but not his expenditure on care. If care is also subject to this constraint (for example, if it involves a dollar investment in safety equipment), then the injurer may have an incentive to invest in *too much* care (Beard 1990). The reason for this paradoxical result is that greater spending on care before an accident reduces the injurer's asset level, which makes it more likely that he will be bankrupt (and hence shielded from liability) in the event of an accident. Thus, from the injurer's perspective, each additional dollar spent on care up front costs less than one dollar, which creates an incentive to spend more.

3.7 The Impact of Liability Insurance

Most individuals who engage in risky activities purchase liability insurance to cover, at least partially, any damages that they may cause to themselves or others. In fact, most states require drivers to purchase accident insurance before they will issue a vehicle registration. Most drivers would purchase insurance willingly, however, because they are risk averse; that is, they are willing to pay some amount of money to avoid random fluctuations in their wealth.[33] (Our analysis to this point has assumed that people are risk neutral.)

The problem with insurance is that it potentially reduces the ability of tort liability to create incentives for care.[34]

In terms of incentives, insurance has a similar effect as the judgment-proof problem by shielding the insured party from some or all of the damages that he or she causes. Although the injurer paid a premium to purchase the insurance in the first place, the premium is a sunk cost at the time of the care choice. Thus, insurance will cause the injurer to take too little care from a social perspective. This moral hazard problem ultimately hurts the injurer, however, because insurance companies are aware of the problem and set the premium up front to reflect the actual risk.

Insurance companies have ways of mitigating moral hazard, however. One is to condition the premium, to the extent possible, on the risk-reducing behavior of insured parties. For example, insurance companies give discounts to those who maintain a good driving record and charge more to those who buy sports cars. Another response is to offer partial coverage. Most insurance policies include deductibles requiring the insured to pay some fixed amount before the insurance kicks in. The higher the deductible, the greater is the injurer's incentive to take care, and hence the lower is the premium. (Raising the deductible is like increasing α in [2.10].) The problem, however, is that this reduces the value of insurance in reducing risk. Optimal insurance coverage thus strikes an efficient balance between risk reduction and incentives (Shavell 1979).

3.8 Litigation Costs

Our discussion of accident costs in this chapter has occasionally mentioned administrative costs as one component of overall accident costs, but we have not explicitly considered how the costs of using the courts to resolve accident claims affects the operation of the tort system. In Chapter 8 we examine how legal costs affect the manner in which parties resolve legal disputes in general (for example, whether to settle or go to trial).[35] Here we examine the prior question of how these costs affect the incentives for injurer care, which determines the number of disputes that arise in the first place.

Generally speaking, litigation costs tend to reduce incentives for injurer care (Hylton 1990; Ordover 1978). To see how, consider the simple unilateral care model with strict liability. Observe first that when litigation is costly, the *efficient* level of care is higher compared to the model with zero litigation costs. This is true because the costs of an accident now include both the damages to the victim and the litigation costs of both the injurer and the victim. Thus, the marginal benefit of care increases while the marginal cost remains the same.

In the model with litigation costs, strict liability gives inadequate incentives for injurer care for two reasons. First, litigation costs will deter some victims from filing suit. Thus, injurers will not face liability for the full amount of damages that they cause (the logic is the same as for the judgment-proof problem). Second, the injurer will ignore the litigation costs incurred by those victims who do file suit.

The outcome may be better under a negligence rule. In this case, if the injurer complies with the due standard of care, victims will be deterred from filing suit because they will not expect to recover any damages at trial. Thus, the zero-litigation-cost outcome is possible because no litigation costs are actually incurred. In reality, however, some victims do file suit under negligence, either because some injurers are in fact negligent, or because the court makes errors, as discussed in the next section. If the due standard of care continues to be set by the Hand rule without concern for litigation costs, then there will be underdeterrence in this case as well because injurers will have no incentive to take more care than is necessary to avoid liability.

3.9 Legal Error *

We have assumed to this point that courts implement the negligence rule without error. This requires both correct measurement of the injurer's actual level of care and correct calculation of the due standard. In reality, however, evidentiary uncertainty will cause courts to make two types of errors in applying the negligence rule: they will sometimes find an injurer nonnegligent when he actually violated the due standard (a type I error), and they will sometimes find an injurer negligent when he complied with the due standard (a type II error). It turns out that both types of errors reduce the incentives for care under negligence (Png 1986).

To see how, let q_1 be the probability of a type I error, and let q_2 be the probability of a type II error. Consider first an injurer who takes care. Under a perfectly functioning negligence rule, he would face no liability, but with error, he faces expected liability of q_2D (where D is the victim's damage). Similarly, an injurer who does not meet the due standard would face liability of D under a perfectly functioning negligence rule, but with error he only faces expected liability of $(1 - q_1)D$. Thus, the injurer's gain from taking care (that is, his savings in liability) is given by

$$(1 - q_1)D - q_2D = (1 - q_1 - q_2)D. \tag{2.12}$$

Note that this expression is less than the gain from taking care in the absence of error, which is D. Further, it is clear from (2.12) that an increase in either type of error reduces the gain from taking care.

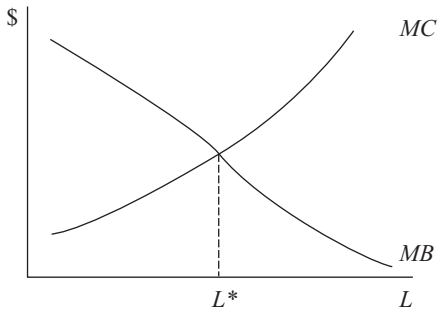

Figure 2.8
The Optimal Statute of Limitations for Tort Suits

3.10 The Statute of Limitations for Tort Suits

A statute of limitations sets an upper bound on the period of time following an accident during which the victim can bring a legal action for compensation. There are two sorts of benefits associated with limiting suits in this way. The first has to do with limiting legal error, which presumably increases with the length of time between the accident and the trial due to fading memories and deteriorating evidence. The second is simply the savings in litigation costs because fewer suits are permitted (Baker and Miceli 2000; Miceli 2000).

Offsetting these benefits are the costs of a limited filing period, which consist of the reduced incentives for injurer care. In particular, as the statute length is shortened, injurers expect to face fewer suits and therefore have less incentive to take care. (A shorter statute therefore has the same effect as a lower value of α in [2.10].) The optimal statute length, L^*, therefore occurs where the marginal benefit of increasing L equals the marginal cost. This is shown graphically in Figure 2.8.

3.11 Intentional Torts

The focus of our analysis to this point has been on *accidental* harms, but there is also an area of tort law concerned with harms that are intentionally caused, such as assault and battery. In this section, we examine the economics of these so-called intentional torts.[36]

We begin by distinguishing two possible meanings of *intentional harm*. The first is harm that is the inevitable consequence of certain risky activities. In other words, the probability of harm from engaging in these activities is nearly certain, even though the activity itself is not meant to cause harm. An example is the manufacture and sale of a dangerous product like a chainsaw. The probability of harm from any single chainsaw might be relatively small, but if enough are sold, the probability approaches one.[37] The second meaning of intent is harm resulting from a single act whose primary objective is to

cause an injury, like throwing a punch. In both of these cases, intent is reflected by a high probability of harm, but intuition tells us that they are different. The question is whether they should therefore be treated differently by the law.

From an economic perspective, one key difference between the two cases is that the injurer in the latter case (the one throwing the punch) is not investing in precaution to avoid an accident but is in fact making an effort to *increase* the probability of harm. Presumably he does this because he derives some benefit from inflicting the harm. Although one might claim that the chainsaw manufacturer is also deriving benefit from selling a product that it knows to be dangerous, the fundamental difference is that it is socially desirable for the manufacturer to invest in a positive amount of accident avoidance, whereas it is optimal for the puncher to invest zero effort to cause harm (assuming that the harm to the victim exceeds the benefit to the injurer).

Despite this difference, strict liability for compensatory damages would seem to achieve the proper level of deterrence in both cases since it forces injurers to internalize the full costs of their actions. In fact, punitive damages are probably called for in the second case for at least two reasons. First, intent to inflict harm implies forethought on the part of the injurer with the likely consequence that he will seek to avoid responsibility. (This is a second key difference between the two notions of intent described above.) As we have seen, when the probability of detection is less than one, punitive damages are needed to achieve optimal deterrence.[38]

Second, in some cases the benefit to the injurer of inflicting harm may exceed the cost to the victim, but the benefit is not socially valuable. Examples are acts of violence like rape and murder. In these cases, compensatory damages are insufficient to deter the injurer, so some additional sanction is needed. This example, combined with our discussion of the judgment-proof problem above, also suggests why intentional torts are sometimes classified as crimes. If the injurer has insufficient wealth to pay damages, then the threat of imprisonment can provide the only deterrent. In Chapter 9 we provide further economic reasons for the use of criminal law, as opposed to (or in addition to) tort law, for the control of certain harmful acts.

3.12 Valuing Human Life and Safety

We conclude this introductory chapter on tort law with a few words on the problem of measuring damages. As we have noted, one of the functions of tort law is to compensate victims for their losses. This is straightforward when the damages are to property, which has a market-determined value, but it is more difficult in the case of personal injury or death. Since compensation necessarily takes the form of monetary damages, however, the court must

place a dollar value on a victim's injuries or loss of life. Typically, courts seek to provide compensation for financial losses as well as for nonmonetary factors such as pain and suffering.[39]

Financial losses are ordinarily based on the present value of lost earnings of the victim, plus medical expenses (if any). In computing lost earnings, one takes account of the victim's educational attainment, life expectancy, fringe benefits, and the like, as well as the amount they would have spent on their own consumption plus the taxes they would have owed on their earnings (if the award itself is not taxed). The job of calculating this amount usually falls to an economist acting as an expert witness for the plaintiff.

Courts also award monetary damages for pain and suffering, though here the calculation is more subjective (and hence more controversial). A key factor in determining the amount of pain and suffering is the impact that the accident has on the victim's loved ones. The greater their perceived loss, the larger the award. Another rationale for awarding pain and suffering is to serve as an additional deterrent when compensation for lost earnings seems inadequate, such as when the victim dies or when the injurer's actions were intentional. In this sense, pain and suffering and punitive damage awards serve similar economic functions.

The lost-earnings approach to valuing personal injury is practical, but it most likely underestimates total costs. For example, it ignores the value of leisure. Another approach is to survey people to find out how much they are willing to pay to avoid certain risks. Such "contingent valuation" surveys are useful when market information is lacking—thus, they are often used in valuing environmental damage, like loss of an endangered species. An interesting issue that arises out of surveys is the gap between willingness-to-pay (WP) and willingness-to-accept (WA) measures of value. It turns out that most people require more compensation to take on an increase in risk (WA) than they will pay to reduce risk by the same amount (WP).

There is, however, a potentially more reliable way than surveys to get information about how people value risks to themselves—namely, wage premiums for risky occupations. Specifically, how much more do workers need to receive to work as, say, miners compared to safer jobs? The resulting measure, referred to as "hedonic damages," reflects the willingness-to-accept measure of risk. Although such studies are useful, they too are subject to bias. First, they only reflect people's valuation of small risks and cannot reliably be extrapolated to obtain a measure of the value of life. (For example, if a worker accepts a wage premium of $100 in return for a 1/10,000 risk of death, does that mean that he values his life at $1 million?) Second, people who accept risky jobs are ordinarily those who incur the lowest cost from the risk, or who may actually receive a benefit from it (for example, firemen and policemen). This will bias the risk-premium downward. Finally, as we will argue in the

next chapter, people often misperceive risk, thereby distorting the ability of the market to adequately compensate workers for it. For these reasons, courts generally look skeptically on hedonic damage measures.

4 Conclusion

This chapter laid out the basic economic theory of tort law. At the center of the theory is the model of precaution, which prescribes that injurers and victims should invest in accident-reducing activities up to the point where the last dollar spent on care equals the marginal savings in accident costs. The role of the law is to provide incentives in the form of liability rules for the parties to meet this standard. We argued that the law of negligence, as embodied by the Hand test, conforms well to this ideal.

We also examined several factors that complicate the simple model. We first considered causation, which plays a prominent role in the law but fits somewhat uneasily into the economic model. We nevertheless proposed several economic explanations for the practical importance of causation. In addition, we showed that activity levels can have a significant effect on the risk of accidents but are not handled well by negligence rules.

Finally, we considered several departures of real-world accidents from the simple model—including legal error, injurer bankruptcy, the availability of liability insurance, and litigation costs—all of which tend to reduce the efficiency of the tort system. We will encounter several of these problems again in the next chapter in the course of applying the economic model to specific areas of tort law.

Discussion Questions

1. We have argued in this chapter that the primary economic function of tort law is to deter unreasonably dangerous behavior, but a second function is to provide social insurance against accidental harm. In what ways are these objectives compatible, and in what ways are they incompatible?
2. Suppose that an accident has occurred. Which of the following does the victim *not* have to prove in order to recover damages under tort law: (1) that she sustained some harm; (2) that the injurer's actions were cause-in-fact of the harm; (3) that the injurer's actions were proximate cause of the harm; (4) that the injurer intentionally caused the harm?
3. State whether a rule of strict liability or negligence is preferred in each of the following situations:

(a) The court makes errors in measuring the level of damages from an accident.
(b) The court makes errors in measuring the injurer's care level.
(c) The victim's care is an important determinant of accident risk.
(d) The injurer's activity level is an important determinant of accident risk.

4. (a) Describe the test courts use to determine cause-in-fact.
   (b) Suppose that a train traveling at 35 mph collides with a car that was stalled at a crossing and injures the driver. The court determines that trains traveling faster than 25 mph are negligent, but it also determines that even if the train in question had been traveling at this slower speed, it could not have avoided hitting the car. Would the train's negligence be found cause-in-fact of the accident? Explain why or why not.
5. What is the economic rationale for a statute of limitations for tort suits? Based on this rationale, explain why it makes sense that there is no statute of limitations for serious crimes like murder.

PROBLEMS

1. Consider a unilateral care accident model in which the injurer can either take care at a cost of $50, or no care. Further, suppose that if he takes care, there is zero risk of an accident, but if he does not take care, the risk is 0.1. Finally, suppose the victim's damages in the event of an accident are $750.
   (a) Use the Hand rule to determine if care is efficient in this case.
   (b) Suppose the victim cannot observe the injurer's care choice and therefore would be unable to prove negligence in the event of an accident. Would it be appropriate for the court to invoke the doctrine of *res ipsa loquitur* in this case? Explain why or why not.
2. A train passing a farmer's property emits sparks that sometimes set fire to the farmer's crops. The crop damage can be reduced, however, if the railroad installs spark arresters on its trains, if the farmer moves his crops, or both. The following table summarizes the cost of the various possible actions, and the crop damage (if any):

Action	*Crop damage ($)*	*Farmer's cost ($)*	*RR's cost ($)*
No action	150	0	0
Farmer moves crops	90	15	0
RR installs arresters	40	0	50
Farmer moves crops and RR installs arresters	0	15	50

(a) Which action yields the socially optimal outcome?
(b) What action will result under a rule of no liability? (Assume here and in subsequent questions that bargaining costs between the RR and farmer are high.)
(c) What action will result under a rule of strict liability?
(d) What action will result under a negligence rule where the due standard of care for the RR is to install arresters?

3. Consider the owner of a large tract of undeveloped land that is suitable for recreational use by campers, bikers, hikers, and so on. If the land is left open, users will enter and enjoy a benefit of $3,000, but they may suffer accidents. Specifically, if users enter and take no care, they will suffer expected damages of $2,000, whereas if they enter and take care of $200, they will suffer expected damages of $500. The owner of the land cannot take care to reduce the risk of accidents, but he can fence in the land at a cost of $1,000, which prevents entry altogether.
   (a) Calculate the net benefit from each of the following options: (1) owner fences land; (2) owner does not fence land, users enter and take no care; (3) owner does not fence land, users enter and take care. Which is most efficient?
   (b) Many states have passed laws that immunize owners of undeveloped land from liability if they open the land for recreational use. What outcome in (a) will result under this rule?
4. The case of *New York Cent. R. Co. v. Thompson* (21 N.E.2d 625, 1939) concerned a woman who accidentally caught her foot in some railroad tracks and was injured when a train failed to brake in time and struck her. Suppose the liability rule is negligence with a defense of contributory negligence.
   (a) Will the woman be able to recover her damages if the court determines that she was negligent for walking on the tracks?
   (b) Under what doctrine will the woman be able to recover, despite her negligence, if the court determines that the train saw her predicament and had time to brake but did not? Explain the economic rationale for this doctrine.
5. Suppose that an injurer escapes liability in three out of four accidents that he causes. Let the average damages per accident be $100,000 and assume the liability rule is strict liability.
   (a) In order for the injurer to face the correct incentives to take care, what should his total damages be in each case where he is held liable?
   (b) What portion of total damages is compensatory and what portion is punitive?

CHAPTER 3

Applying the Economic Model of Tort Law

The preceding chapter focused on general principles of accident law that apply to a broad range of accident settings. In this chapter we apply these principles to several specific areas of tort law. We begin with products liability, or accidents caused by dangerous products. We devote the most attention to this topic, both because it has become an important area of tort law and the source of much dissatisfaction with the operation of the tort system, but also because it raises some new conceptual issues for the economic theory of accident law, the primary one being the distinction between accidents involving "strangers" and accidents involving parties to a contract. Following our discussion of products liability, we examine (in a more cursory fashion) workplace accidents, environmental hazards, and medical malpractice.

1 Products Liability

The number of products liability suits increased markedly during the decades of the 1980s and 1990s. To get an idea of the numbers involved, look at Table 3.1, which shows the number of products liability cases filed in U.S. District Courts from 1980 to 1998. Despite some fluctuations during the 1990s, the trend is predominantly upward, even when measured as a percentage of all civil cases filed during this period (see Figure 3.1).[1] One result has been an increase in the price of certain consumer products, another the withdrawal of some from the market altogether. These trends have led to a number of proposals for tort reform, some of which we discussed in the preceding chapter (for example, the call for a cap on punitive damage awards).

TABLE **3.1** **Data on Products Liability Cases Filed in U.S. District Courts, 1980–1998**

Year	*Total cases filed*	*Cases as a % of all civil cases*
1980	6,876	4.07
1981	8,028	4.45
1982	7,908	3.84
1983	8,026	3.32
1984	7,677	2.94
1985	12,507	4.57
1986	12,459	4.89
1987	14,145	5.92
1988	16,166	6.75
1989	13,408	5.74
1990	18,679	8.57
1991	12,399	5.97
1992	10,769	4.75
1993	16,545	7.24
1994	23,977	10.16
1995	17,631	7.38
1996	38,170	14.00
1997	23,294	8.79
1998	28,325	10.84

SOURCE: *Viscusi (1991, tables 2.1, 2.2);* Statistical Abstract of the U.S., *various years.*

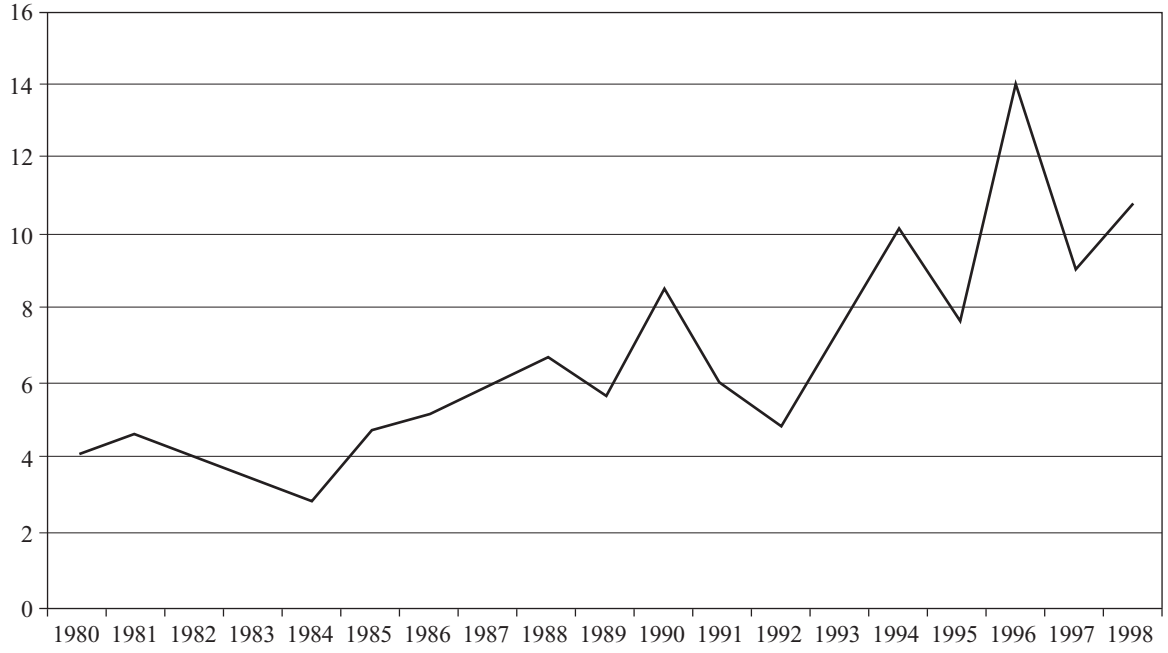

Figure 3.1
Products Liability Cases as a Percentage of All Civil Cases Filed in U.S. District Courts, 1980–1998

We begin our discussion of product-related accidents in this chapter by briefly reviewing the history of products liability law in the United States. We then extend the model from the previous chapter to show the impact of the market relationship between the injurer and the victim. Our objective is to explain the historical trends in terms of the economic model and to evaluate the current status of the law.

1.1 A Brief History of Products Liability Law

In contrast to the modern image of products liability law as protecting defenseless consumers against manufacturers of dangerous products, the law in the nineteenth century was based on the belief that excessive producer liability would burden society with high administrative costs and threaten the economic viability of business.[2] The past 150 years, however, have witnessed a gradual evolution in the law in the direction of greater producer liability. This has occurred in several distinct phases.

The first phase began in the mid-nineteenth century with the birth of the doctrine of "privity," which held that in the event of a product-related accident, the purchaser only had a cause of action against the immediate seller of the product—that is, the party with whom he had a direct contractual relationship.[3] For example, if an automobile accident occurred as a result of negligence on the part of the manufacturer, the victim could only seek recovery from the retailer.

Under privity, the allocation of risk from product-related accidents largely relied on contract rather than tort principles. Although we will see below that the chain of contractual relationships leading from the manufacturer to the ultimate consumer can theoretically serve to shift liability from immediate sellers back to the manufacturer, in reality, this shifting occurs imperfectly. Thus, the doctrine of privity effectively insulated most manufacturers from liability.

The privity limitation nevertheless endured through the end of the nineteenth century until it was finally overturned in 1916 in the famous case of *MacPherson v. Buick* (217 N.Y. 382, 111 N.E. 1050, 1916). The case involved an accident that occurred when one of the wheels on the plaintiff's car broke off, causing him to be thrown from the car. Since the plaintiff had bought the car from a dealer, the doctrine of privity apparently barred the plaintiff from recovering against the manufacturer. Judge Benjamin Cardozo rejected this position, however, based on the argument that the manufacturer could clearly have foreseen the possibility of injuries to individuals other than the immediate purchaser of the car (in this case, the dealer). This did not immediately imply liability on the part of the manufacturer, however. The victim also had to prove negligence by the manufacturer (which he succeeded in doing in *MacPherson*). Nevertheless, the transition from no liability to negli-

gence had occurred, thereby significantly expanding the scope of producer liability.

The next phase in the evolution of products liability law, which witnessed the transition from negligence to strict liability, occurred by two separate routes. The first was the result of a gradual increase in the standard of care owed by product manufacturers and sellers. A key case in this development was *Escola v. Coca-Cola Bottling Co.* (24 Cal.2d 453, 150 P.2d 436, 1944), which concerned an injury caused by an exploding Coke bottle. Although the plaintiff, who was a waitress in a restaurant, could offer no evidence of negligence on the part of the manufacturer, the court held the manufacturer liable based on the doctrine of *res ipsa loquitur*. Recall that under this doctrine, the fact of the accident itself is evidence of negligence—only a defective Coke bottle would explode. As noted in the previous chapter, the application of *res ipsa loquitur* in cases where due care does not entirely eliminate the risk of accidents amounts to a rule of strict liability.

The second route to strict liability occurred in the area of producer liability for breach of warranty. Under the theory of warranties, a branch of contract law, sellers were strictly liable for damages caused by products that failed to function as represented—considerations of negligence were irrelevant. However, the requirement of privity remained for these cases because warranties (implied or expressed) are a form of contract.

This changed with the 1960 case of *Henningsen v. Bloomfield Motors, Inc.* (32 N.J. 358, 161 A.2d 69, 1960). The case also concerned an automobile accident, this time caused by a failure of the steering mechanism. The new element of this case was that the sale contract between the plaintiff's husband and the manufacturer included a clause that expressly limited the latter's liability to the original purchaser and for only certain types of damages. The court rejected this type of contractual limitation, however, arguing that the implied warranty of fitness prevailed regardless of any expressed intentions of the parties to the contrary. Further, the court struck down the privity requirement, noting that, although the victim was not the purchaser, she was someone who "in the reasonable contemplation of the parties to the warranty, might be expected to become a user of the automobile. Accordingly, her lack of privity does not stand in the way of prosecution of the injury suit against the defendant Chrysler."

With the *Henningson* decision, the tort and contract theories of products liability had converged on a strict liability standard. This was explicitly recognized with the publication of the Restatement (Second) of Torts in 1965, Section 402A of which says:

(1) One who sells any product in a defective condition unreasonably dangerous to the user or consumer or to his property is subject to liability

for physical harm thereby caused to the ultimate user or consumer, or to his property, if
(a) the seller is engaged in the business of selling such a product, and
(b) it is expected to and does reach the user or consumer without substantial change in the condition in which it is sold.

(2) The rule stated in Subsection (1) applies although
(a) the seller has exercised all possible care in the preparation and sale of the product, and
(b) the user or consumer has not bought the product from or entered into any contractual relation with the seller.

Note that part (2)(a) excludes consideration of producer care (hence, liability is strict), while part (2)(b) eliminates privity.

To say that liability is strict, however, is somewhat misleading because, in addition to causation, plaintiffs must show that the product had a *defective design*, or, if it is *inherently dangerous* (like cigarettes or dynamite), that the manufacturer failed to warn consumers of the danger. Thus, there is an element of negligence in strict products liability because manufacturers can avoid liability by meeting the design standard or the duty to warn. Recent trends, however, have made it harder to meet these standards.

With the foregoing history as background, in the next section we develop a formal model of products liability with the objective of explaining the broad trend in the law toward strict producer liability. The crucial extension in the accident model from the previous chapter will be to explicitly account for the contractual relationship between the injurer (producer) and the victim (buyer).[4]

1.2 An Economic Model of Products Liability

We develop our analysis of product-related accidents in the context of a simple model of perfect competition.[5] As a benchmark, we first consider a product for which there is no risk of an accident. Let the aggregate inverse demand curve for the "safe" product be given by $b(q)$, which represents the amount consumers are willing to pay for a unit of the good as a function of the number of units purchased, q. A downward-sloping demand curve (reflecting diminishing marginal benefits) implies that $b(q)$ is decreasing in q, as shown in Figure 3.2.

On the supply side, we assume, for simplicity, that marginal and average costs are constant and equal to c.[6] Thus, the supply curve is horizontal at c. Equilibrium output for the safe product occurs at the point where demand equals supply, or at q^* in Figure 3.2, while the equilibrium price is equal to

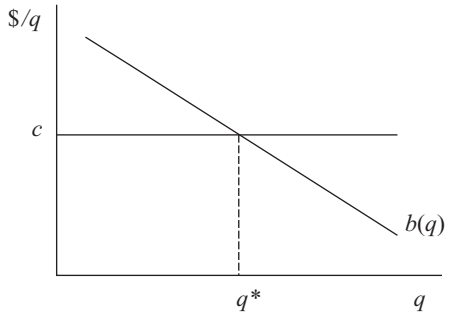

Figure 3.2
Equilibrium Output and Price for a Safe Product

the constant marginal cost, c. Algebraically, equilibrium output is defined by the equation $b(q^*) = c$.

1.2.1 Equilibrium Price and Output for a Dangerous Product

Now consider a product for which there is a risk of injury to the consumer, but which is identical to the safe product in all other respects. Assume that each unit of the product carries the same probability of an accident, p, and the same damages in the event of an accident, D. Thus, total expected damages are qpD (for now we suppress considerations of care, or the safety of the product). Note that this specification mirrors our discussion of activity levels in Chapter 2, where we assumed that expected damages were proportional to the injurer's (or victim's) activity level.

The existence of accident risk may affect either the demand or the supply sides of the market (or both), depending on how the law assigns liability between the manufacturer and the consumer. For purposes of the current discussion, we represent the liability rule as follows. Let s represent the share of accident costs borne by the manufacturer and $1 - s$ the share borne by the consumer, where s is between zero and one. Note that all of the rules from Chapter 2 emerge as special cases of this general formulation. For example, $s = 1$ corresponds to strict liability, $s = 0$ corresponds to no liability (also known as caveat emptor, or "buyer beware"), and conditioning s on the injurer's and/or the victim's care level can yield the various negligence rules.

Consider first the impact of the accident risk on the demand side of the market. In comparison to the safe product, we would expect consumers to reduce their willingness to pay for a unit of a risky product by exactly the amount of their expected accident losses. Thus, if consumers will pay $b(q)$ for a unit of the safe product, they will pay $b(q) - (1 - s)pD$ for a unit of the risky product, where $(1 - s)pD$ is the uncompensated portion of their expected damages. This has the effect of shifting the demand curve down relative to that in Figure 3.2. This is illustrated in Figure 3.3 for the case of strict

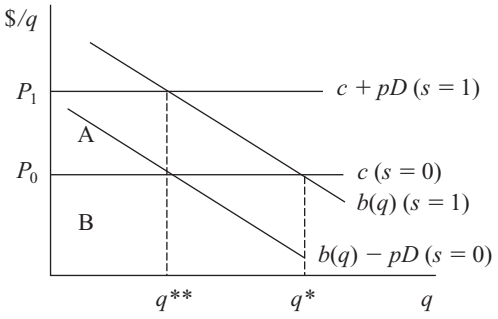

Figure 3.3
Equilibrium Output and Price for a Dangerous Product Under Different Liability Rules

liability ($s = 1$) and no liability ($s = 0$). Note that the demand curve for the risky product is equivalent to that for a safe product when the rule is strict liability because the consumer expects to be fully compensated in the event of an accident, but the curve shifts down by the full expected damages under a rule of no liability because the consumer expects to receive no compensation.

Now consider the impact of accident risk on the supply side. The marginal cost, or supply curve in this case will equal marginal production costs plus expected liability per unit of output, or $c + spD$. Thus, under a rule of no liability ($s = 0$), the supply curve corresponds to that for the safe product, but under strict liability ($s = 1$), the supply curve shifts up by the full amount of expected damages. These two curves are also shown in Figure 3.3.

As before, equilibrium output and price are determined by the intersection of the relevant demand and supply curves. Figure 3.3 shows the equilibrium under strict liability and no liability. The first thing to note about these two equilibria is that *they result in the same level of output, q*** (which is less than the equilibrium output of the safe product, *q**). This is not a coincidence. In fact the result can be stated more generally: *equilibrium output in the model with accident risk is independent of the liability rule*. To prove this general result, equate the algebraic expressions for demand and supply to get

$$b(q) - (1 - s)pD = c + spD. \tag{3.1}$$

But note that terms multiplied by s on the left- and right-hand sides cancel, leaving the condition for equilibrium output:

$$b(q^{**}) = c + pD \tag{3.2}$$

which is independent of s. Thus, no matter how the law assigns liability for accidents, equilibrium output occurs at the point where marginal consumption benefits for the good ($b[q]$) equals the *total* marginal costs, including marginal production and accident costs.

To see the intuition for this result, we first need to determine the equilibrium price. Note in Figure 3.3 that, unlike output, price is *not* independent of

the liability rule. In particular, under strict liability ($s = 1$), the price is given by $P_1 = c + pD$, the full marginal costs (including expected accident costs), whereas under no liability ($s = 0$), the price is $P_0 = c$, or simply marginal production costs. The difference reflects the fact that, under strict liability, the manufacturer is selling the consumer the product *bundled with an insurance policy for the associated accident risk*. Thus, the price reflects the marginal production costs (c) plus the expected damages (pD), where the latter in effect acts like an actuarially fair insurance premium. In Figure 3.3, the insurance component of the price is therefore the difference between P_1 and P_0, while area A is the aggregate expected damages that the manufacturer expects to pay out. Algebraically, area A is given by the insurance premium, pD, multiplied by the aggregate output, q^{**}.

In contrast, when the rule is no liability (caveat emptor), the price simply reflects the marginal production costs because the manufacturer faces no liability in the event of an accident. Area B thus equals aggregate production costs. Consumers nevertheless must still pay for the expected damages, but now they expect to pay it out of their own pockets when an accident occurs. This is what causes the demand curve to shift down when $s = 0$, with the result that the equilibrium output remains at q^{**}. Although consumers cannot look to manufacturers to insure them against product risk in this case, most will have purchased some form of health insurance that will cover any damages due to product-related accidents. The discounted price for the product provides funds that can be used to purchase this insurance.

Alternatively, consumers can "self-insure" by setting aside an amount pD for every unit of the dangerous product that they purchase. Over the long run, this will provide exactly enough money to compensate them for their losses (again, area A in Figure 3.3). The problem with this approach is that if the first unit purchased results in an accident, the consumer will not have had time to accumulate enough resources. This is one important reason why consumers are better off purchasing market insurance rather than self-insuring. Firms, especially small ones, are susceptible to this same problem, so under a rule of strict liability they too usually purchase market insurance to cover their expected tort liability.

EXERCISE 3.1

Let the aggregate inverse demand curve for a dangerous product be given by $b(q) = 20 - q$. Also, let

$c = \$5$

$p = .01$

$D = \$1{,}000.$

Derive the equilibrium output and price for the product under a rule of no liability ($s = 0$) and under a rule of strict liability ($s = 1$).

The discussion of products liability to this point, and in particular the fact that equilibrium output does not depend on the liability rule, represents an example of the Coase Theorem, discussed in Chapter 1. Recall that the Coase Theorem says that when parties to a legal dispute can bargain at low cost, they will allocate resources efficiently regardless of the particular assignment of liability. As Figure 3.3 and equation (3.2) show, the equilibrium level of output for a dangerous product occurs at the point where marginal consumption benefits equal total marginal costs, regardless of the liability rule. Output is thus invariant to the assignment of liability. The reason for this is the shifting of liability by means of the price.

As we will see below, however, when the price mechanism fails to function perfectly, the requirements for the Coase Theorem are no longer satisfied, and the liability rule will matter for efficiency. This was the case in the model of accidents between "strangers" in the previous chapter. "Strangers" in this sense means parties who had no contractual or market relationship prior to the accident. As a result, they had no opportunity to bargain over the allocation of liability, or at least the cost of doing so was prohibitively high. (When you get into your car, imagine the prospect of bargaining with all the motorists or pedestrians with whom you might have an accident.) In that case, the liability rule was crucial in determining the allocation of resources. Indeed, recall that our discussion of activity levels in the previous chapter ended with the conclusion that, for accidents between strangers, none of the standard liability rules could achieve efficiency of care and activity levels by both injurers and victims.

1.2.2 Care Choices by Manufacturers and Consumers

In focusing on equilibrium output, we have to this point ignored the care choices of the manufacturer and consumers. The question in this context is whether the irrelevance of the liability rule extends to care as well. In theory, the answer is yes, again as a result of the Coase Theorem. To see why, suppose initially that the rule is no liability. In the model of accidents between strangers, we saw that victims will take efficient care under this rule, but injurers will take no care. In the product model, however, suppose that the manufacturer and consumer can strike a bargain whereby the manufacturer agrees to produce a safe product in return for a higher price to reflect the extra cost.[7] If this bargaining exhausts all gains from trade, then the manufacturer will invest in safety to the point where the marginal reduction in accident risk just

equals the marginal cost—in other words, he will invest in the efficient safety level.

A similar story holds for strict liability. Under this rule, the problem with the stranger model was too little victim care. To remedy this inefficiency, the necessary bargain would entail a promise by the consumer to use the product carefully in return for a price reduction by the manufacturer. Again, if the bargain exhausts all gains from trade, it will yield the efficiency level of victim care.

The preceding shows that the Coase Theorem holds for care as well as output, assuming that the market mechanism functions perfectly. In assessing whether these bargains will actually occur, however, the reader may have perceived an asymmetry in the two cases. Under no liability, the consumer will pay the higher price provided that she perceives that the product is indeed safer. In many cases, this will merely require careful inspection of the product prior to purchase. (In other cases, the increased safety will have to be taken on faith or may be misperceived, as discussed below.)

In contrast, the bargain under strict liability requires that the consumer must honor her promise to use the product carefully *after she has paid the lower price and taken possession*. Given the cost of care, this creates a situation in which the consumer may renege on her promise with little if any chance of detection by the producer. As a result, the producer is unlikely to be willing to engage in the proposed bargain in the first place. Such a "market failure" undermines the Coase Theorem in this case.

The preceding discussion suggests that a pure strict liability rule will probably not achieve efficiency regarding consumer care. That still leaves several candidates for a fully efficient rule, including no liability, the various negligence rules, and strict liability with contributory negligence. It turns out, however, that we can significantly narrow the list by considering another possible source of market failure—consumer misperceptions of risk.

1.2.3 Consumer Perceptions of Risk

Our discussion of the economic model of accidents, whether involving strangers or market participants, has assumed that the parties correctly estimate the risk of an accident. In fact, individuals tend to *misperceive* risk in a systematic way. Specifically, there is evidence that they tend to overestimate low probability risks and underestimate high probability risks (Viscusi 1991, 64). Since product accidents primarily fall in the former category (low probability risks), we would expect consumers to systematically overestimate them.

Let us consider how consumer misperception of risk, whether over- or underestimation, affects the analysis of products liability law (Spence 1977; Polinsky and Rogerson 1983). To begin, we initially ignore care and focus on

the determination of equilibrium output. Suppose that consumers potentially misperceive the probability of an accident, viewing it to be αp rather than p, where $\alpha > 1$ represents an overestimate, and $\alpha < 1$ represents an underestimate. Assume, however, that producers perceive p correctly, reflecting the superior knowledge they have about their product's risk, and assume that both consumers and producers correctly estimate the damages from an accident, D.

In this setting, the demand curve for an arbitrary legal rule becomes $b(q) - (1 - s)\alpha pD$, while the supply curve remains the same as before, $c + spD$. Equating these expressions yields the condition defining equilibrium output:

$$b(q) - (1 - s)\alpha pD = c + spD. \tag{3.3}$$

Note first that when $\alpha = 1$, (3.3) is identical to (3.1)—this is the case of no misperceptions where output is independent of the liability rule. However, when $\alpha \neq 1$, (3.3) and (3.1) will differ for any liability rule other than strict liability (that is, for any $s < 1$), meaning that equilibrium output in these cases will depart from the efficient level of output, q^{**}, depending on the specific liability rule and the nature of consumers' perceptions.

To illustrate, suppose that the rule is no liability ($s = 0$), which corresponds to the lower pair of supply and demand curves in Figure 3.3. Setting $s = 0$ in (3.3) yields

$$b(q) - \alpha pD = c. \tag{3.4}$$

If consumers overestimate risk, $\alpha > 1$, and the demand curve is below that in Figure 3.3. As a result, output is too low—consumers demand too little of the product. In contrast, if consumers underestimate risk, $\alpha < 1$, and the demand curve is above that in Figure 3.3. Output in this case is too high. This conclusion generalizes to the case of any $s < 1$; so long as consumers expect to bear some of their own losses, misperceptions will affect output in the directions just described, though the extent of the inefficiency decreases as s approaches one.

EXERCISE 3.2

Reconsider the example from Exercise 3.1, but now suppose that consumers misperceive risk. Assume that the rule is no liability ($s = 0$).

(a) Calculate the equilibrium output when consumers overestimate the risk to be .012 (rather than its true value of .01).
(b) Calculate the equilibrium output when consumers underestimate the risk to be .008.

The conclusion is different when the producer is strictly liable. In that case, we have seen that (3.3) reduces to (3.1) and output is efficient for any α. Misperceptions have no effect on output in this case because the consumer internalizes the expected damages through the market price, which accurately reflects the risk, given our assumption that producers have no misperceptions.

What this discussion shows is that when consumers misperceive risk, the liability rule matters for efficiency of the output level. The general conclusion is that *the party who more accurately perceives the risk should bear the liability in equilibrium*. This argument supports the historical trend toward strict liability in conjunction with the increasing complexity of most consumer products.

We now reintroduce care into the analysis. We saw above that strict liability induces efficient producer care but will not provide incentives for consumer care due to the high cost of enforcing contracts conditioned on consumer use of the product. Incentives for consumer care can be restored, however, by including a contributory negligence defense as discussed in the previous chapter.

1.2.4 A Note on Custom as a Defense

If a particular safety feature becomes widespread in an industry, it may achieve the status of "custom." The existence of industry custom provides a potential standard for applying a negligence rule in products liability cases.[8] For example, showing a manufacturer's failure to adhere to industry customs is an easy way for consumers to prove negligence, and courts have historically accepted such arguments. The question, however, is whether adherence to custom should be accepted as evidence that the manufacturer is *not* negligent.

The most famous custom case, *The T.J. Hooper* (60 F.2d 737, 2d. Cir. 1932) explicitly addressed this question. (Interestingly, the opinion was written by Judge Learned Hand of the *Carroll Towing* case.) The case concerned a tugboat that lost its barge and cargo during a coastal storm because it did not have a radio by which the captain could have been warned. The owner of the barge sued claiming negligence, but the tug owner argued that it was not customary for tugs to have radios, and hence his failure to have one was not negligence. Judge Hand rejected this argument, stating that

> in most cases reasonable prudence is in fact common prudence; but strictly it is never its measure; a whole calling may have unduly lagged in the adoption of new and available devices. It may never set its own tests, however persuasive be its uses. Courts must in the end say what is required; there are precautions so imperative that even their universal disregard will not excuse their use.

Judge Hand thus argued that cost-benefit principles trump custom in determining negligence. But why would market forces not compel an industry to adopt all cost-justified safety measures? (That is, why wouldn't the Hand test and custom arrive at the same solution?) The answer is consumer misperceptions of risk, which limit the ability of the market to enforce efficient safety standards. For this reason, custom is properly rejected by courts as a defense in products liability cases. (See the further discussion of this issue in the context of medical malpractice in Section 4.1 below.)

1.2.5 Recent Trends

Our conclusion to this point is that strict liability with contributory negligence is the most efficient liability rule for product-related accidents. Strict liability provides incentives for manufacturers to produce safe products and ensures that the market price accurately reflects the residual risk to consumers, while the contributory negligence defense gives consumers an incentive to use the product properly.

The general trend in the law toward strict liability during the first half of the twentieth century seems consistent with this conclusion. However, recent developments in the law appear to be stripping producers of some defenses against liability while holding them liable for some risks that were unforeseen or unknowable at the time of manufacture of the product. (An example is the risk from asbestos—see Section 3.3 below.) Some commentators refer to this emerging standard as "absolute" or "enterprise" liability. The economic model suggests that this could reduce efficiency for those products where consumer misuse is an important determinant of accidents (Priest 1988).

1.2.6 Evidence on the Impact of Products Liability Laws

Most economic analysis of products liability is theoretical, but a couple of studies have examined the effect of products liability laws on prices. For example, in a study of the market for childhood vaccines, Manning (1994) found that wholesale prices for several vaccines have increased dramatically in the past few decades as a result of increasing producer liability. Further, a substantial portion of this increase has been due to litigation costs. Manning (1997) similarly found a liability premium in the cost of prescription drugs in the United States as compared to Canada, reflecting the significantly higher liability costs in this country. Although these studies confirm the prediction of a higher product price in response to greater producer liability, they cannot tell us whether consumers have received their money's worth in terms of safer products, and/or more efficient insurance against risk.

1.3 Concluding Remarks

We conclude the discussion of products liability by emphasizing the reasons why contract law is not an adequate remedy for most product-related accidents. Although the injurer and victim have a contractual relationship, which we have seen can fully internalize the accident risks, we have also seen that various sources of market failure can inhibit this mechanism from functioning perfectly. These include consumer misperceptions of product risk and the inability of producers to monitor consumer use of the product. A further reason is simply the cost of writing contract terms in the presence of remote risks, which, as Landes and Posner (1987, 281) observe, "may well be disproportionate to the benefit of a negotiated (as distinct from imposed-by-law) level of safety."

These conclusions illustrate the general principle, asserted in Chapter 1, regarding the role of the law in internalizing costs. In particular, when bargaining between the concerned parties can occur smoothly, the specific legal rules do not matter for efficiency—the primary role of the law is to enforce whatever contracts the parties write. This is the insight of the Coase Theorem. However, when bargaining fails, the law needs to be more interventionist in assigning liability. According to economic theory, this is where contract law needs to give way to tort law. The history of products liability law in the twentieth century seems to provide an example of this transition, though some would argue that recent developments have caused it to overshoot the mark.

2 Workplace Accidents

This section deals with accidents in the workplace, including accidents in which workers are injured on the job as a result of unsafe working conditions or negligence by a fellow worker, as well as accidents in which a worker causes an injury to a nonworker (a stranger) in the course of his or her employment. Many of the issues raised by the first type of accident—those in which the victim is a worker and the injurer is the employer or another employee—have already been discussed in our analysis of products liability. For example, in a perfectly functioning labor market, the wage will adjust to reflect the legal assignment of liability between the parties. Thus, contract rather than tort law principles can, in principle, govern these accidents, though market failures of the sort discussed above may again interfere with the attainment of an efficient outcome. In contrast, accidents involving a worker and a nonemployee raise many of the same issues discussed in Chapter 2 in the

model of accidents between strangers. The emphasis in this section will therefore be on unique aspects of the law governing workplace accidents.

2.1 Respondeat Superior

Under the doctrine of respondeat superior, an employer is strictly liable for accidents caused by his employees' negligence when committed in the course of their employment.[9] One possible rationale for this rule is that employees will often lack the resources necessary to compensate victims' losses (that is, they will be judgment proof). The law therefore allows victims to reach into the "deep pockets" of their employers. While this "vicarious liability" of employers makes sense regarding the compensation function of tort law, it may be an impediment to efficient accident avoidance since it insulates the injurer from responsibility for damages.

It is possible, however, that the employer can use his contractual relationship with the employee to give the latter an incentive to be careful. For example, the employer can supervise employees and threaten to fire those who perform their duties in a careless manner. As Landes and Posner (1987, 121) note, "Making the employer liable for his employee's tort serves to enlist the employer as a substitute enforcer of tort law where the primary enforcement mechanism, a tort action against the immediate tortfeasor is unworkable."

2.2 Accidents in which the Victim Is an Employee

An important exception to the liability of employers for their employees' negligence concerns accidents in which the victim is also an employee. Historically, employer liability for these accidents was severely limited. The common law did impose a duty on employers to maintain a safe workplace and to warn of dangerous situations, but even an employer who was negligent in fulfilling these duties could defend himself by demonstrating contributory negligence or assumption of risk by the injured worker. As we saw in the previous chapter, a rule of negligence with a defense of contributory negligence provides efficient bilateral incentives for care, and, as we saw in the case of products liability, the wage will adjust to compensate workers for whatever losses they cannot recover from the tort system (as well as to achieve the efficient level of employment).[10]

A further defense was available to employers when an employee was injured as a result of the negligence of another employee. Although the doctrine of respondeat superior would seem to have imposed strict liability on the employer in this case, the so-called fellow servant rule actually absolved the employer of any liability, provided that the latter had not been negligent in hiring or inadequately supervising the negligent employee. The economic ra-

tionale for this rule is that it gives workers an incentive to monitor one another and to report careless behavior to the employer.

While the fellow servant rule might have been appropriate in small enterprises and shops where workers had close contact with one another, it seems less valid in large businesses where victims might be injured by the negligence of workers with whom they had never had contact (Keeton et al. 1984, 571). More cynical observers simply saw the rule as yet another pro-business rule that, like the privity limitation in products liability, insulated firms from liability. For whatever reason, the law governing workplace accidents changed dramatically in the early twentieth century.

2.3 Workers' Compensation Laws

Following the turn of the twentieth century, dissatisfaction with the common-law rules governing workplace accidents led to legislation by all states that instituted a form of strict employer liability. Employer negligence was no longer necessary for recovery, nor could the employer invoke contributory negligence or the fellow servant rule as defenses. The new laws differed from strict liability, however, in that the amount of compensation was set by fixed damage schedules for each class of injury. (A typical formula calls for replacement of two-thirds of wages for a set period of time.) In addition, agencies rather than the courts administered the rules.

In evaluating the efficiency of these laws, we can draw an analogy to products liability, where we argued that, although the price mechanism can in principle shift risks in such a way as to make the particular liability rule irrelevant (according to the Coase Theorem), market imperfections like misperceptions of risk make this mechanism unreliable in practice. In this setting, we argued that strict liability imposes the risk on the party who can best estimate it (the firm), and the wage or price can adjust appropriately.

A possible inefficiency in workers' compensation laws is the elimination of contributory negligence as a defense, which may result in too little care by workers. This problem may not be severe, however, for two reasons (Landes and Posner 1987, 310–11). First, employers can contract with workers to achieve the efficient level of safety by paying a higher wage for greater care. Second, the limitation on compensation mitigates the moral hazard problem associated with standard strict liability.

To see the latter point, suppose that in the event of an accident, a worker expects to receive fixed compensation equal to $\overline{D}$, while her actual damages would be $D(x, y)$, where, recall, x is the employer's care and y is the worker's care. The worker's choice of care will therefore solve

$$\text{minimize} \quad y + p(x, y)[D(x, y) - \overline{D}]. \tag{3.5}$$

It is possible to show that the victim will choose more care than under true strict liability (which, in this simple model, results in zero victim care), though she will choose less than optimal care. The victim has an incentive to take some care at the margin because by doing so, she reduces the amount of undercompensation.

Another inadequacy in workers' compensation laws is that the victim must prove that the injury is job-related. This is straightforward when the injury is the result of an accident, but for illnesses like cancer that have multiple causes, the burden is more difficult. The problem is one of "uncertainty over causation" as discussed in Chapter 2. Although this potentially attenuates the incentives for employers to provide a safe workplace, recall that efficient incentives can be maintained under two rules. The first imposes full liability on the employer if the conditional probability that the illness is work-related exceeds a threshold, and the second imposes liability on the employer in proportion to the conditional probability that the illness is work-related.

A further check on workplace safety is direct regulation by OSHA, the Occupational Safety and Health Administration. Established in 1970, this agency's goal is to assure "safe and healthful working conditions" for all workers. The most favorable evidence available, however, suggests that it has had only limited success in this effort. Specifically, Viscusi found that over the period 1973–83 OSHA regulations did not significantly reduce work-related injuries and illnesses, and they reduced lost workdays by only 1.5–3.6 percent (Viscusi 1986). One explanation for this is the high cost of monitoring compliance. Another may simply be that the threat of liability for workers' compensation had already given employers an incentive to take most cost-justified safety measures, so further improvements in safety were hard to come by.

3 Liability for Environmental Damages

This section discusses issues that arise in the use of tort law for internalizing environmental damages. The role of tort law in this context is generally limited to unanticipated releases of harmful substances like oil spills or toxic-waste leaks, referred to as "environmental accidents." In contrast, the continuous discharge of pollutants as the known by-product of a firm's production process is usually dealt with by means of Pigovian taxes or direct regulation. We will examine these regulatory approaches to environmental policy in Chapter 7 as part of a general discussion of the control of externalities.

3.1 Characteristics of Environmental Accidents

Environmental accidents are similar in many ways to other sorts of accidents, but they also present some unique problems. This section emphasizes those unique elements.

3.1.1 Multiple Victims

Many environmental accidents involve multiple victims. Examples include radiation discharges from nuclear power plants and oil spills. One problem created by the existence of multiple victims is that, while aggregate damages may be large, the damage suffered by any individual victim may be too small to justify the cost of filing suit against the injurer. This is referred to as the *dispersed cost problem*. To illustrate, suppose that n victims each suffer individual damages of D dollars, making aggregate damages nD. Also, let the cost to any one victim of filing suit be c dollars. If $D < c$, no victim will find it *privately* worthwhile to file suit, even though a suit is *socially* desirable, given that $nD > c$.

One solution to this problem is a class-action suit in which all of the individual claims are bundled into one suit. This not only overcomes the disincentive of individual victims to bring suit, but also it economizes on judicial resources by eliminating duplicative trials over the same set of factual and legal issues. In most cases, these benefits will more than offset the costs of identifying and notifying all victims (underinclusion), as well as preventing uninjured parties from claiming to be victims (overinclusion).

A second problem associated with multiple victims is that the likelihood of injurer bankruptcy increases. Suppose that the injurer has total assets of A out of which it can pay liability judgments. In the previous example, the injurer will not be able to cover all damages if $A < nD$, a situation that becomes more likely as n increases. Not only does this result in undercompensation of victims, we also saw in the previous chapter that it potentially reduces incentives for injurer care, depending on the liability rule. When the rule is strict liability, the possibility of insufficient assets generally reduces the incentive for injurers to take care because their expected liability is less than the full damages that they impose.[11] (In particular, the injurer expects to pay liability of A dollars when damages are $nD > A$ dollars.) In contrast, under a negligence rule the injurer may still have an incentive to take efficient care because by doing so he avoids *all* liability. Thus, if the savings in liability from choosing due care, equal to A dollars, is larger than the cost of taking the additional care, then the injurer will do so.

The preceding suggests that a move toward strict liability may have the un-

intended effect of creating incentives for firms engaging in hazardous activities to alter their organizational structure so as to use bankruptcy as a shield against liability. There is in fact evidence that firms engage in this sort of strategy by contracting out particularly hazardous aspects of their business into smaller firms (Ringleb and Wiggins 1990).

3.1.2 Causal Uncertainty

A second distinguishing feature of environmental accidents is that the particular cause of the accident may not be easy to identify. One circumstance in which this causal uncertainty arises is when there are *multiple injurers*. For example, several polluters may dump hazardous waste into a landfill, which eventually seeps into the groundwater. Another example, not in an environmental context, is when two hunters fire at what they think is a deer, and one of their bullets hits a third hunter.[12] This situation, in which the actions of multiple injurers contribute to a single harm, is sometimes referred to as a *joint tort*.[13]

To illustrate the problems created by joint torts, consider the following example of two injurers whose actions create a risk of damages to a single victim. Let $p(x_1, x_2)$ be the probability of an accident as a function of the expenditures on care by the two injurers, and let D be the fixed damages in the event of an accident. Note that this resembles our model of bilateral care except that now it is the two injurers who take care rather than the injurer and the victim. As before, the social problem is to choose the care levels to

$$\text{minimize} \quad x_1 + x_2 + p(x_1, x_2)D. \tag{3.6}$$

By analogy to the bilateral care model, optimal care by both injurers in this case requires that each face the victim's full damages at the margin. In general, this will not be possible given the constraint that the total liability payments collected from the injurers cannot exceed the damages suffered by the victim. To illustrate, suppose that the rule is strict liability and that each injurer is responsible for a share of total damages. Specifically, suppose injurer one pays a share s_1 and injurer two pays a share s_2 where $s_1 + s_2 = 1$. The problem facing each injurer is therefore to

$$\text{minimize} \quad x_j + p(x_1, x_2)s_jD, \quad j = 1, 2. \tag{3.7}$$

Like the judgment-proof problem, both injurers face less than full damages (that is, $s_j < 1$) and therefore take too little care. (Compare the problem in [3.7] to that in [2.10] in the previous chapter.)

How are the shares determined in actual law? The traditional common-law rule is that the victim can collect her full damages from either injurer or from

both. In the latter case, the victim can obtain the judgment in whatever proportions she chooses, usually based on which injurer is best able to pay. Under this rule, each injurer must form an expectation about his share of damages, but the constraint that the shares must sum to one implies that neither injurer will generally expect to face full damages. As a result, they will both tend to take too little care.

The conclusion is different under a negligence rule. In this case, it is possible to show that if the due standard of care is set at the efficient level for each injurer, then in equilibrium they will both meet the standard. The intuition is the same as in our earlier discussions of the negligence—each injurer has an incentive to meet the due standard in order to avoid any share of the victim's damages.

A second source of causal uncertainty is when there is a *long latency period* between the exposure to a toxic substance and the emergence of the illness. The problem here is that when the illness emerges, there may be no way to tell whether it was due to the accidental exposure or to a normal "background" or "natural" risk.[14] This is the situation we examined in the previous chapter under the heading of "uncertainty over causation." We showed there that efficient incentives for injurer care can be achieved by using ordinary strict liability or negligence rules with no limitation on liability to reflect the background risk.

We also showed that it is possible to maintain efficiency if liability is limited in one of two ways. The first is a threshold rule that holds the injurer liable only if the conditional probability that he caused the accident exceeds an appropriately chosen threshold. The second is a rule that holds the injurer liable for the proportion of the damages that he caused in a probabilistic sense, conditional on the fact that the illness actually occurred.

All of the preceding rules assign liability only after an exposure victim has actually contracted the illness. Another approach to causal uncertainty is to allow all victims of the exposure to file for damages *at the time of exposure*. In this case, the risk is itself at tort (a "tort for risk"), and damages are calculated to reflect reduced life expectancy, future pain and suffering, and future medical costs resulting from the exposure.

To see how the proportional liability and tort-for-risk rules compare, suppose that damages from an illness, when contracted, are $150,000; the probability of getting the illness from the accidental exposure is .10; and the background probability is .05. The overall probability of developing the illness after exposure is therefore .15. Under the proportional liability rule, the share of damages the injurer pays equals the conditional probability that the exposure caused the illness, given that the illness has occurred. This probability is $(.10)/(.15) = 2/3$. Thus, at the time the illness occurs, the injurer pays $(2/3)(\$150{,}000) = \$100{,}000$. In contrast, under the tort-for-risk, the injurer

would pay damages *at the time of exposure* equal to the contribution of the exposure to the expected losses from the illness, or (.15 − .05)($150,000) = $15,000.

Although it may seem that the tort-for-risk rule imposes less cost on the injurer and therefore provides less incentive for care, recall that proportional damages are not paid to all victims, but only those who develop the illness. Thus, the injurer's expected cost under the proportional rule, as of the time of exposure, is (.15)($100,000) = $15,000. The two rules therefore provide identical (and efficient) incentives. The rules are not identical in all respects, however. The chief advantage of the proportional rule is that it saves on litigation costs since not all exposure victims end up filing suit. The advantage of the tort-for-risk rule is that it avoids the risk that the injurer will be judgment-proof at the time, possibly well in the future, when the illness occurs.

3.2 Superfund

An important area of environmental law concerns the cleanup of hazardous waste sites.[15] Despite the obvious risk to public health and the environment from these sites, there was little regulatory oversight of disposal practices prior to the 1970s. An important change occurred in 1980 with the enactment of the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA). The primary objective of this legislation was to clean up hazardous waste sites quickly and effectively and to impose the cost (when possible) on the responsible parties. To fund the cleanup of sites, CERCLA established a "superfund" to be financed in part by taxes but also by damage actions brought by the Environmental Protection Agency (EPA) against responsible parties. It is the liability aspect of CERCLA that is of interest to us here.

The extent of liability under CERCLA is broad. First of all, liability is strict, and in the case of multiple polluters, it is "joint and several." This means that any one of them can be held responsible for the entire cost of cleanup. Thus, "disposal of a thimbleful of hazardous waste at a large disposal site exposes an entity to enormous potential liability" (Menell 1991, 109). The resulting uncertainty has led to dramatically higher costs of insurance for environmental liability, when it is available at all.

In addition to holding polluters strictly liable, CERCLA extends liability to "innocent" buyers of a contaminated site. Many have criticized this provision as discouraging transactions that would otherwise lead to the beneficial redevelopment of old industrial sites—so-called brownfields. This negative conclusion is not necessarily true. Recall from our discussion of products liability that, in the absence of misperceptions about risk, the equilibrium output of a dangerous product is independent of the allocation of liability between the buyer and the seller. The same is true here; if land prices accurately

reflect anticipated cleanup costs, then there should be no distortions in land transactions. However, if sellers have better information about the extent of contamination than do buyers, then too few transactions may occur as a result of adverse selection (Segerson 1997).

To illustrate, suppose that a contaminated site is worth V dollars to a buyer (developer) and R dollars to a seller, exclusive of cleanup costs, which equal C. Since someone must pay the costs whether or not a sale occurs, it is efficient for the buyer to acquire the site if $V > R$, which we assume is true. Let s be the share of costs that the buyer expects to incur. If P is the price, the buyer will purchase the site if

$$V - sC \geq P. \tag{3.8}$$

As for the seller, if no sale is made, she must pay the full cleanup costs (assuming she is solvent), yielding a value of $R - C$, whereas if she sells, she receives the price less her share of cleanup costs, or $P - (1 - s)C$. She will therefore sell if $P - (1 - s)C \geq R - C$, or if

$$P \geq R - sC. \tag{3.9}$$

A sale will occur if there exists a price that satisfies both (3.8) and (3.9); that is, if $V - sC \geq R - sC$, or if $V > R$, which is the condition for an efficient sale. This shows that, regardless of s, the efficient outcome will occur.

It should be easy to see, however, that if the parties hold different assessments of the size of C, this conclusion will no longer hold. Suppose, in particular, that the seller has a better assessment of C due to private information. In that case, the efficient transaction will only occur if the seller bears full liability (that is, if $s = 0$) since the buyer's criterion in (3.8) will be independent of C (except insofar as it is reflected in the price). Note that this conclusion mirrors the above argument that strict products liability is efficient because it imposes liability on the party with better information about risks (the seller). However, it appears contrary to the imposition of liability for environmental contamination on "innocent" buyers ($s = 1$).

Under the original provisions of CERCLA, lenders could also be held liable for cleanup if the owner was insolvent (Segerson 1993). Again, if credit markets operate perfectly, this creates no distortions and in fact helps to mitigate the incentive problems due to insolvency of the injurer. However, if there is asymmetric information between the borrower (injurer) and lender, an adverse selection problem of the sort described above arises. In addition, if the injurer makes some or all of its abatement (care) decisions *after* the loan is made, and the terms of the loan cannot be made contingent on the level of care (for the same reason that the seller of a dangerous product cannot condition the price on the buyer's care after purchase), then the injurer will have an incentive to underinvest in abatement.

The discussion in this section has pointed to several problems in use of tort liability and its statutory counterpart for internalizing environmental harm. In some cases, these problems can be overcome by redesigning liability rules in ways that we have discussed, but in others, a liability approach is inherently flawed. As a result, an efficiently designed approach to the control of environmental externalities will likely involve a combination of liability and safety regulation, a topic to which we return in Chapter 7.

3.3 Case Study: Asbestos

Asbestos is a product that was widely used in the United States in a variety of industrial settings, as well as in schools and homes. The link between exposure to asbestos and severe illnesses like lung cancer, asbestosis, and mesothelioma, however, apparently was not known until the 1930s. Tort suits against asbestos manufacturers began in the 1970s, slowly at first, but by the 1990s, they had expanded to the point where they comprised a substantial fraction of all products liability cases filed in this country.

Asbestos litigation involves several of the problems that we have identified with the use of the tort system for internalizing risk (Dewees 1998). These problems largely stem from the long latency period of asbestos-related illnesses, usually ten to thirty years. First, it is difficult for plaintiffs to prove causation given the existence of multiple background risks unrelated to exposure. Second, plaintiffs may be unable to establish which of several manufacturers or suppliers was responsible for their exposure. And even if they could, the responsible party may have gone bankrupt by the time the illness arises. As we have seen, these factors weaken the deterrence function of tort law.

Asbestos is also an interesting case study because it combines aspects of products liability, workplace safety, and environmental risk. Initially, workers' compensation provided the sole remedy for work-related exposures to asbestos, but since compensation is limited, plaintiffs' lawyers early on sought to circumvent that system. They succeeded in 1973 by suing manufacturers (rather than employers) under products liability principles.[16] This success resulted in a surge of tort claims in the 1970s that has continued into the 1990s. (See Figure 3.4, which shows the trend in asbestos cases filed in U.S. District Courts from 1974 to 1998.) Evidence that manufacturers knew of the risks of asbestos and failed to warn workers in many cases resulted in punitive damage judgments against defendants. The resulting financial pressure caused several bankruptcies, the most notable being that of the largest manufacturer, the Manville Corporation, in 1982.

Today, the risk of new asbestos exposures has been greatly reduced, partly as a result of this litigation. One could therefore argue that the law has suc-

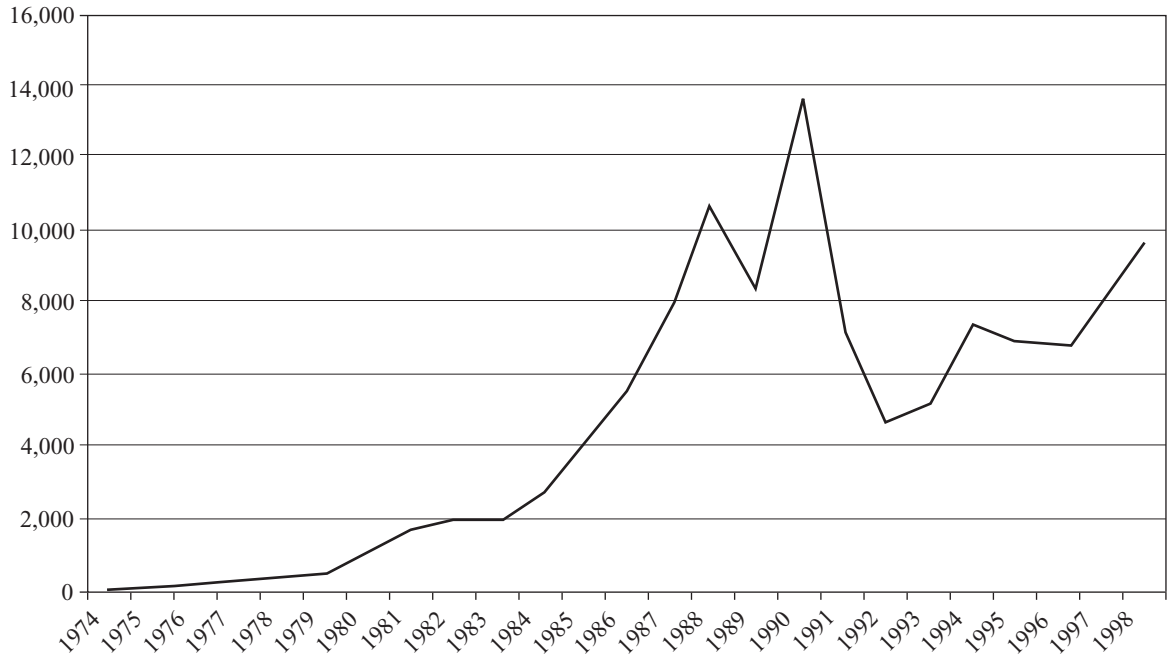

Figure 3.4
Asbestos Cases Filed in U.S. District Courts, 1974–1998

ceeded in its deterrence function. Most would agree, however, that it has been much less successful in compensating victims, while at the same time imposing high costs on the legal system.

One reform that could improve the compensation function of tort law in mass-exposure cases would be to allow victims to file at exposure for expected damages, rather than having to wait until actual symptoms arise. The advantage is that all victims would receive some compensation, which they could use to purchase health insurance or precautionary medical treatment. The drawback is the likely "flood" of litigation (Robinson 1985). To date, some states have taken the limited step of allowing exposure victims to collect medical monitoring expenses, but none has gone so far as the allow a full-blown "tort for exposure."[17]

4 Medical Malpractice

Following the trend of other forms of tort litigation, patient claims against physicians for malpractice have risen significantly in recent decades. Begin-

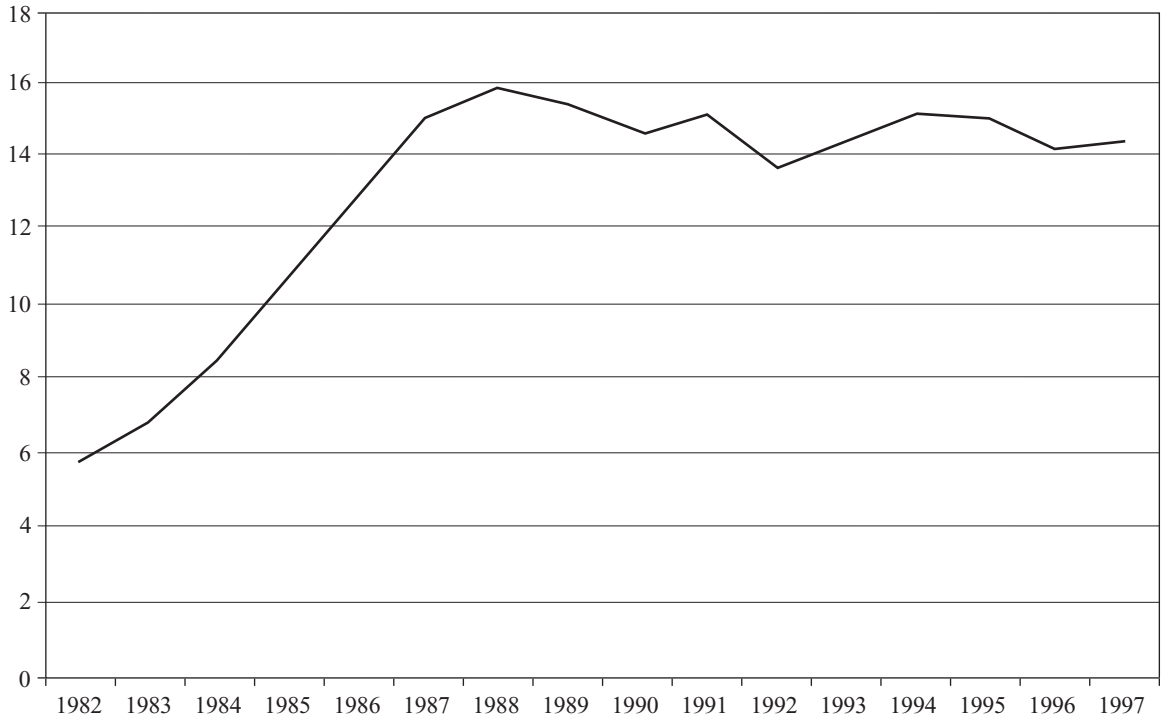

Figure 3.5
Mean Liability Premium ($1,000s), All Physicians, 1982–1997

ning in the late 1960s and continuing through the 1970s and 1980s, malpractice claims rose at roughly 10 percent per year, while damage awards and malpractice insurance costs correspondingly increased (Danzon 1991). (See Figure 3.5, which shows the trend in malpractice insurance premiums.) Though some of this increase can be attributed to wider availability and use of certain higher risk medical procedures, the general trend mirrors the growth in products liability and other tort claims over the same period. In response to this "medical malpractice crisis," many states enacted reforms, including caps on awards, a shorter statute of limitations for malpractice claims, and regulation of legal fees (Keeton et al. 1984, 192–93).

The economic analysis of medical malpractice resembles products liability and workplace accidents in the sense that the injurer and victim have a preexisting market relationship. Thus, contractual principles can theoretically govern the assignment of liability. In most cases, however, consumers of medical care, especially high-risk care like surgery, are infrequent purchasers. Further, they lack the knowledge to evaluate the quality and/or safety of care. The likely misperception of risk therefore prevents an efficient contractual solution for reasons noted above.

Another source of inefficiency that precludes a contractual solution in this area is insurance—the majority of consumers have either private or public health insurance that pays most of their medical expenses. Because the price consumers face is therefore "too low" (and the provision of care is not otherwise rationed), the consumption of medical services is distorted, apart from considerations of risk.

4.1 Customary Practice and Informed Consent

The actual liability rule applied to malpractice cases is negligence, with the due standard based on "customary practice." Specifically, "the doctor must have and use the knowledge, skill and care ordinarily possessed and employed by members of the profession in good standing" (Keeton et al. 1984, 187). Thus, a successful finding of negligence generally requires expert testimony by another physician. Some criticize this standard as giving the medical profession (and other professions subject to the same standard) the right to establish their own standards of conduct (especially given the reluctance of physicians to testify against one another). Further, we argued above that custom as a defense is unlikely to establish an efficient standard when consumers are poorly informed about available technologies. One difference here is that physicians have a code of professional ethics to provide the best possible care for their patients, a factor that may outweigh the profit motive.

A more recent duty imposed on physicians also helps to mitigate the problem of asymmetric information between doctors and patients. It is the duty of doctors to inform patients of the risks involved in receiving a particular treatment—the *doctrine of informed consent*. Disclosure is costly, however, so the law only compels physicians to inform the patient of "material risks," defined as risks that a reasonable person would find significant. This reflects an effort by courts to efficiently balance the costs and benefits of information.

4.2 Do Physicians Practice Defensive Medicine?

The increasing costs of malpractice litigation have led many to ask whether physicians practice "defensive medicine." According to Danzon (1991, 54), "Defensive medicine should be defined as liability-induced changes in medical practice that entail costs in excess of benefits and that should not have occurred in the absence of liability." Generally speaking, one might interpret defensive medicine as excessive care by physicians ($x > x^*$) in response to an expectation of liability in excess of actual damages imposed. A difficulty in measuring the extent of such activities is that the moral hazard problem associated with consumer health insurance can also result in too much care from the demand side, as noted above.

The logical limit of defensive medicine is that many physicians simply stop performing high-risk procedures. This reflects an inefficiently low activity level, assuming that the benefit of the activity exceeds the risk. Unfortunately, the only evidence on the extent to which physicians alter their practices (too much care and/or too little activity) in response to the threat of liability comes from surveys and anecdotal evidence, neither of which is a good basis for drawing general conclusions or for formulating public policy.

Continuing dissatisfaction with the malpractice system has led to renewed calls for reform. One interesting alternative is to institute a no-fault system along the lines of workers' compensation. Under this proposal, strict liability would replace negligence, and adjudication would be taken out of the tort system and transferred to an administrative agency. The intended benefits would be reduced delay in the litigation process, better compensation of victims, and greater predictability of physicians' liability. In addition, it is consistent with our conclusion above that strict liability is preferred to negligence when consumers misperceive risk.

5 Conclusion

This chapter applied the economic model of accidents to products liability, workers' compensation, environmental accidents, and medical malpractice. While these areas of tort law differ in many respects, we showed that economic theory can go a long way toward explaining their basic features.

An important methodological issue that arose in our discussion of products liability (but applied to other areas as well) was the role of the contractual relationship between injurers and victims in internalizing accident risk. We showed, in particular, that when parties correctly perceive risk and can bargain with one another, the output of the product will be efficient regardless of the liability rule; only the equilibrium price will adjust to reflect the assignment of liability. This was in contrast to accidents between "strangers" (parties who cannot bargain), or where the market failed to operate efficiently, in which case the liability rule matters for efficiency. This distinction between disputes where the parties can bargain and those where they cannot will be a recurrent theme in the chapters ahead.

Discussion Questions

1. Discuss the difference between accidents between strangers and accidents in which the injurer and victim have a preexisting economic (contractual) relationship.