

AN ANALYSIS OF CAUSATION AND THE SCOPE OF LIABILITY IN THE LAW OF TORTS

STEVEN SHAVELL*

I. INTRODUCTION

ONE person harms another. For the victim to succeed in a suit in tort against the injurer, he must establish two things. He must show that the accident¹ in which he suffered harm is within the scope of liability—that it is not of a type which certain legal principles, notably “causal” in nature, make ineligible for further consideration. And, assuming the accident is within the scope of liability, he must also demonstrate that the injurer is obligated to pay him damages under an applicable rule of liability.² It is not enough for him to show only the second of these two elements. For example, under the negligence rule, it is not enough for the victim to show that the injurer was negligent. Suppose that smoke suffocates a sleeping man in his bed at a hotel that had failed to install a proper fire escape.³ This accident would be found outside the scope of liability on the principle that the absence of a proper fire escape was not a “but-for” cause or “cause in fact” or condition *sine qua non* of the harm: the man would still have suffocated had there been a proper fire escape. Or suppose that an adult thoughtlessly hands a loaded gun to someone’s young child who then drops it on his toe, thereby suffering an injury;⁴ or that a streetcar traveling at excessive speed happens to be at just the point

* Assistant Professor, Harvard Law School. I wish to thank Guido Calabresi, Peter Diamond, Robert Ellickson, Michael Graetz, Jerry Green, Bruce Greenwald, James Krier, William Landes, A. Mitchell Polinsky, Richard Posner, Robert Rosenthal, and Henry Steiner for comments and the National Science Foundation (Grant No. SOC 78-20159) and the Liberty Fund for financial support.

¹ “Accident” will refer generically to an instance in which harm is done. This usage should not be misleading since for the most part attention will be limited to torts that are not intentional in nature.

² Of course, the notion that the application of a rule of liability is distinct from the determination of the scope of liability is to some degree a convenience for organizing thought; in practice, the consideration of the two may be merged. However, see note 41 *infra*.

³ *Weeks v. McNulty*, 101 Tenn. 495, 48 S.W. 809 (1898); and like cases noted in William L. Prosser, *Law of Torts* 238 (4th ed. 1971).

⁴ Second Restatement of Torts, Section 281, Illustration 3.

on its route at just the "right" time to be hit from above by a falling tree.⁵ Despite the negligence of the injurers in these two cases, the accidents would probably be seen as falling outside the scope of liability on the principle⁶ that the way the harm came to pass was in some sense coincidental to the negligent act, that it bore "no relation" to the act, or that it was not "within the normal risk" created by the act. Suppose that after drinking nitroglycerine from an unguarded bucket left at the mouth of a mine a dog stumbles and explodes, injuring a person nearby.⁷ This accident might well be judged to be outside the scope of liability on the ground that it was a freak occurrence, that leaving the bucket of nitroglycerine was not the "proximate cause" of the harm, that the dog was an "unforeseeable intervening cause," or that the harm was not a "natural and probable consequence" of the negligent act. Again, the victim would not be able to collect damages from the negligent party.

What purposes are served by restricting the scope of liability in such cases as these? On what arguments can we coherently explain as socially advantageous the causal and other legal requirements which are generally used to limit the scope of liability?⁸ This is the question to be addressed here.⁹ In answering it, I will adopt the "instrumentalist" approach to the analysis of law and consider in a self-conscious and careful way how the law functions to promote postulated social goals, given assumptions about the behavior of individual parties. This was the approach taken by Professor Calabresi in a recent article on causation in tort law, and the present article builds on his.¹⁰

⁵ *Berry v. Borough of Sugar Notch*, 191 Pa. 345, 43 A. 240 (1899).

⁶ Notice that the former principle concerning but-for causation could also be applied in the case of the loaded gun (if the gun had not been loaded, presumably it would still have fallen on the child's toe), whereas it could not be applied under the circumstances of *Berry* (if the streetcar had not been speeding, it would not have been struck from overhead).

⁷ This is a hypothetical based on an accident reported in *Time Magazine*, Dec. 20, 1926, at 22, and cited in Leon Green, *The Rationale of Proximate Cause* 195 (1927).

⁸ References made here to the principles actually employed in restricting the scope of liability will usually be to Anglo-American law, but the contours of the relevant law in other countries today (including the socialist countries) are much the same. See A. M. Honoré, *Causation and Remoteness of Damage*, 11 *Int'l Encyclopedia of Comparative Law*, ch. 7 (1971).

⁹ Analogous questions could also be asked in regard to restricting the extent of liability—the magnitude of damages. Consider, for example, *Polemis (In re Polemis and Furness, Withy & Co., [1921] 3 K.B. 560)*, which illustrates the issue of unforeseeability as to the extent of harm done. In *Polemis* a workman negligently dropped a wooden plank into the hold of a ship and instead of doing minor damage, the plank caused a spark which ignited gasoline vapor and resulted in an explosion that destroyed the ship. Although the extent of liability can be of great practical importance, as it certainly was in this case, it will be sufficient for us to limit discussion of the issue to a brief subsection.

¹⁰ See Guido Calabresi, *Concerning Cause and the Law of Torts*, 43 *U. Chi. L. Rev.* 69 (1975). (Note 77 *infra* briefly comments on the similarities and differences between Calabresi's article and the present one.) However, I also want to acknowledge at the beginning my intellectual debt to Henry Edgerton, *Legal Cause I, II*, 72 *U. Pa. L. Rev.* 211 & 343 (1924), especially to part II.

A highly stylized model of the occurrence of accidents, of individual behavior, and of the legal system will first be considered. In doing so, the following aggregate measure of social welfare will be employed: the benefits parties derive from engaging in their activities less the costs of taking care less expected accident losses and less certain "administrative costs" associated with the legal system. (Thus there is neither concern with the sharing or spreading of risk nor with the distribution of income.)

Given the model and the measure of social welfare, questions about the desirability of restricting the scope of liability can be viewed in a natural way as questions about a comparison of certain social benefits and social costs. Two such benefits will implicitly enter in the analysis. First, restricting the scope of liability results in a decline in the administrative costs connected with the use or threatened use of the legal system. The second benefit is a potential one, and, although relevant under the negligence rule, it is easier to explain when liability is strict. Under strict liability it is not hard to imagine circumstances where a party decides against engaging in an activity when it would have been socially worthwhile for him to have gone ahead. Consider a firm that uses a carcinogenic substance in producing a good that we agree ought to be produced because the benefits to consumers of the good exceed the costs of production plus the costs of an increased incidence of cancer among the firms' employees. Were the firm liable for *all* cases of cancer among its employees, then it might well be forced out of business, for it would be paying not only for the increased incidence of cancer due to its activities, but also for the general incidence of cancer due to such factors as pollution from other sources and medical x-radiation. By appropriately restricting the scope of liability, this type of disadvantageous outcome, to be described as the result of *crushing* liability, can sometimes be avoided.¹¹

Only one social cost of restricting the scope of liability will enter in the analysis, namely, decreased incentives to reduce accident losses. Under strict liability, incentives would be dulled simply because parties who do harm would have to pay damages less frequently. However, under the negligence rule the dulling of incentives would come about mainly, or at least most obviously, for a slightly different reason: the highest due care standard to which parties who might do harm could be motivated to adhere would fall as a consequence of less frequent exposure to liability.

Given these social benefits and costs, it will be argued that the set of legal principles employed to determine the scope of liability is socially rational. That is, *given* the assumptions of the model, it will be *deduced* that the principles of causation in fact, of harm occurring within the normal risk created by an act, etc. implicitly serve to maximize social welfare and thus

¹¹ But this is not meant to deny the desirability of discouraging parties from undertaking those activities for which the true costs exceed the benefits.

ought to be employed to limit the scope of liability. But, of course, the degree to which this statement about the model explains or justifies the use in practice of the legal principles depends on the importance of observed departures from the model; and some of the discussion in the latter parts of this article may help the reader to come to a judgment about this.

To be able to attach unambiguous meaning to statements about causal relationship standard concepts of decision theory are used in Part II to define several notions of causation. In Part III the model of the occurrence of accidents and of the functioning of the legal system is described. In Part IV the ideal or "first-best" solution to the accident problem is considered. The first-best solution is a notional solution. It is what a social authority would order done if it wished to maximize social welfare. Equivalently, the first-best solution is what potential parties to an accident would agree to do if they were able to come together beforehand and conclude a mutually satisfactory bargain. The first-best solution provides insight into the solutions to the accident problem that can be achieved under tort law given the possible choices of the scope of liability and of the rule of liability. In Part V achievable solutions are examined when the liability rule is strict liability, and in Part VI when it is the negligence rule. The analysis in Parts IV-VI is informal—a series of suggestive numerical examples is discussed—ar should therefore be accessible to the widest audience; formal proofs of results are presented in the Appendix. In Part VII the model is applied and, where necessary, extended in a discussion of certain legal principles and theories (namely, unforeseeability, the "risk theory," intervening causes, proximity of causation in time and space) used to determine the scope of liability; and the model is applied in a discussion of several issues (harm caused by several injurers, contributory behavior of victims, sharing of risk, extent of liability) which sometimes bear on or are connected with our subject. In Part VIII the view of causation presented here is compared with those advanced by other writers on tort law. The reader can probably get a good impression of the content of this article by looking at the summaries to Parts V and VI and then reading Parts VII and VIII.

II. DEFINITIONS OF CAUSATION

With regard to the actions of a party, there appear to be two basic notions of causation implicit in normal discourse. The first is retrospective in nature; it applies when a consequence has already occurred but a different consequence would have resulted if the party had altered his action in some specified way. The second notion of causation is a prospective one and applies when the party's action affects the likelihood of the future occurrence of a consequence (or set of consequences).

As noted, the conceptual framework of decision theory will be employed

to define these notions of causation.¹² Decision theory is concerned primarily with the problem faced by a single party who must make a choice under conditions of uncertainty. The theory involves three basic elements: states of the world, consequences, and actions. A *state of the world* is a description of the way the world could be which is so complete that it does not leave out any relevant aspect. A *consequence* is anything that could possibly matter to the decision making party. And an *action* is a decision of the party; accordingly, it determines the consequence that would result conditional on each state of the world that could occur.¹³ Let us illustrate these concepts.

Example 1. A driver must choose between two actions, the consequences of which are as shown in Table 1. Here, the action "speed" results in the consequence "accident" if the state is "roads icy" and it results in the consequence "no accident" if the state is "roads not icy"; and similarly for the action "don't speed."

TABLE 1

ACTIONS	STATES	
	Roads Icy	Roads Not Icy
Speed	accident	no accident
Don't Speed	no accident	no accident

Let us now define the first notion of causation.

DEFINITION. One action is a *cause in fact*¹⁴ of a consequence relative to another action if, given the state of world, the consequence would have been different had the second action been taken.¹⁵

¹² The approach of decision theory to problems of decision making under conditions of uncertainty dates back to Bernoulli in the eighteenth century and is today widely accepted among economists and statisticians; it is an extremely versatile approach with an axiomatic basis that many consider to be intellectually compelling. The best discussion of decision theory of which I am aware is that of one of its important contributors, Leonard J. Savage. See in particular pp. 6-17, 20-21, 27-30, and 56-68 (which do not require of the reader any mathematical or statistical knowledge) of his *The Foundations of Statistics* (2d revised ed. 1972). See also Howard Raiffa, *Decision Analysis* (1968); and two essays of Kenneth J. Arrow, *Alternative Approaches to the Theory of Choice in Risk-Taking Situations and Exposition of the Theory of Choice Under Uncertainty*, both in his *Essays in the Theory of Risk-Bearing* (1971).

¹³ If we denote a state of the world by s and a consequence by c , then an action is formally identified with a function $a(\cdot)$ from the set of states of the world to the set of consequences, where $a(s)$ is the consequence that would result given s and a .

¹⁴ As mentioned before, in reference to tort law the terms *cause in fact*, *but-for cause* (but for the action having been taken, the consequence would not have occurred), and a condition *sine qua non* are used interchangeably, and we shall do the same.

¹⁵ More precisely, suppose that action a is taken, that a' is another action, and that the state of the world is s —so $a(s)$ is the consequence. Then a is a *cause in fact* of the consequence $a(s)$ relative to a' if $a'(s) \neq a(s)$.

To illustrate, refer to Example 1 and suppose that the driver speeds and the roads are icy, so that an accident occurs. Then speeding is a cause in fact of the accident relative to not speeding. However, suppose that the driver does not speed and the roads are not icy, so that no accident occurs. Then not speeding is not a cause in fact of not having an accident relative to speeding.

Two points about the relation between the definition of causation in fact and the meaning of causation in everyday speech should be noted. First, although the definition involves an explicit comparison of consequences under different actions, in practice we often leave the comparison of actions implicit. We may say that going 60 mph caused an accident without saying, relative to going at some speed less than, say, 45 mph. Second, we may sometimes employ a slightly generalized notion of causation in fact, according to which reference is made not to a single state of the world and a single consequence but rather to a set of states of the world and/or a set of consequences. We may say that when the roads were icy, going 60 mph caused an accident and we really mean the following. The roads had ice at least .05 inches thick—corresponding to a set of states of the world—and going 60 mph resulted in some type of accident—corresponding to a set of consequences (and going less than 45 mph would not have resulted in an accident). Points analogous to these two will also apply to the cognate notions of causation discussed below and we will not bother to repeat them.

In order to define the second notion of causation, we must refer to *probabilities* describing the likelihoods of the states of the world. The probabilities will be taken to be “subjective” or personalistic (but since for the most part it would do no harm to instead interpret them as “objective,” some readers may wish to do so).¹⁶

DEFINITION. One action is a *probabilistic cause*¹⁷ of a consequence relative to another action if the probability of occurrence of the consequence is higher given the first action than given the second.¹⁸

¹⁶ The idea of subjective probability has been rigorously developed and refined over the years and is an important feature of decision theory; key references are Savage & Arrow, *supra* note 12; and the articles collected in *Studies in Subjective Probability* (Henry Kyburg & Howard Smokler eds. 1964). For explanations of the relationship between subjective and objective probability, and for an argument why the latter concept is not necessary for use even in scientific context, see especially the discussion of exchangeability in the important article by Bruno de Finetti, *Foresight: Its Logical Laws, Its Subjective Sources*, 7 *Annales de l'Institut Henri Poincaré* (1937), reprinted in *Studies in Subjective Probability*, *supra* at 95-138; and see also Savage, *supra* note 12, at 67.

¹⁷ The notion of probabilistic cause has been a part of most legal or other treatments of causation with a probabilistic orientation. The legal treatments in which it is given great stress are those of the German school which elaborate the theory of “adequate cause.” See particularly the discussion of J. von Kries’s ideas in H. L. A. Hart & A. M. Honoré, *Causation in the Law* 411-39 (1959). In addition, see Calabresi, *supra* note 10. (His causal linkage corresponds to probabilistic cause.) See also the probabilist-logician Patrick Suppes’s monograph, *Probabilistic Theory of Causality* (1970), which includes a valuable bibliography on causation.

¹⁸ Formally, suppose that a and a' are two different actions and that c is a consequence. Then a is a probabilistic cause of c relative to a' if $Pr(s|a(s) = c) > Pr(s|a'(s) = c)$.

In Example 1, if the probability of icy roads is positive, then speeding is a probabilistic cause of an accident relative to not speeding.

The relationship between the two notions of causation is that one action is a probabilistic cause of a consequence relative to another action if and only if it is more likely that the first action will be a cause in fact of the consequence than that the second action will be.¹⁹ It is clear, however, that causation in fact does not imply probabilistic causation; an action may often turn out to be a cause in fact of a consequence without being a probabilistic cause (relative to any other action). This is illustrated by the situation of Example 2 modeled on the facts of *Berry*. (See Table 2.)

TABLE 2

ACTIONS	STATES		
	No Wind Storm	Wind Storm; Tree in Middle of Route Is Blown Over	Wind Storm; Tree at End of Route is Blown Over
<i>Probability</i>	.98	.01	.01
Speed	no accident	no accident	tree hits streetcar
Don't Speed	no accident	tree hits streetcar	no accident

Example 2. A streetcar conductor must choose between two actions. If he speeds, the streetcar will reach the end of the route in time to be hit by the tree if the third state of the world occurs; speed will therefore turn out to be a cause in fact of an accident. If the conductor does not speed, the streetcar will be in the position to be hit if the second state of the world occurs, and not speeding will turn out to be a cause in fact of an accident. Therefore, the probability of an accident is .01 regardless of the action and neither action is a probabilistic cause.

When the consequences are losses (measured in some natural unit), we shall occasionally use a notion that is closely related to probabilistic causation.

DEFINITION. One action is a *cause of expected loss* relative to another if losses are expected to be higher on average given the first action than given the second.²⁰

Although the above definitions refer to actions taken by a single decision maker, they can be generalized to actions of two (or more) parties by holding

¹⁹ This follows because $[Pr(s|a(s) = c) > Pr(s|a'(s) = c)] \Leftrightarrow [Pr(s|a(s) = c \text{ and } a'(s) \neq c) + Pr(s|a(s) = c \text{ and } a'(s) = c) > Pr(s|a(s) \neq c \text{ and } a'(s) = c) + Pr(s|a(s) = c \text{ and } a'(s) = c)] \Leftrightarrow [Pr(s|a(s) = c \text{ and } a'(s) \neq c) > Pr(s|a(s) \neq c \text{ and } a'(s) = c)]$.

²⁰ An action *a* is a cause of expected loss relative to *a'* if $Ea > Ea'$, where *E* gives the mathematical expectation of a random variable.

constant the actions of one of the parties and varying the actions of the other. For instance, given that the second party took his action, one action of the first party is a cause in fact of a consequence relative to another action he might have taken if the consequence would have been different had he taken the other action. However, we will not need to think about such modifications of the definitions of causation until we consider in Part VII the subjects of multiple injurers and of contributory behavior of victims.

We have focused in this part on the idea of an action as a cause, yet when we speak of a cause in ordinary discourse we sometimes mean that a state of the world (or set of states—an event) was the cause. To describe this idea, the definitions of causal relationship may be modified, analogously to what was done in the last paragraph, by holding constant the actions of the party and by varying the state of the world. For example, one state of the world is a cause in fact of a consequence relative to another state if, given the action, the consequence would have been different had the second state occurred. But it should be observed that when we speak of a state of the world as a cause, frequently the state was unlikely; otherwise it would not ordinarily be referred to as a cause but rather as a condition.²¹ The example is given of a man who strikes a match to light his pipe while standing in the middle of a field of dry grass. A sudden and strong gust of wind blows the match from his hand; the match lands on the grass and a fire starts. We would be likely to call the gust of wind a cause of the fire and the presence of oxygen a condition rather than the reverse, even though each was necessary to the occurrence of the fire and even though under our definition each was a cause in fact of the fire (relative to states of the world in which there is no sudden gust of wind and there is no oxygen in the atmosphere, respectively).

III. OUTLINE OF THE MODEL

Accidents involve two types of parties, “injurers” and “victims,” and all accident losses fall on victims in the absence of a liability rule. Moreover, the injurers and victims are assumed to be strangers, or at least not to have a contractual relationship.²²

Injurers have to make two types of decision: whether to engage in a particular activity,²³ and what level of care to exercise if they choose to engage

²¹ This point is made in Hart & Honoré, *supra* note 17, at 33, and the example that follows is taken from the same place.

²² The results would carry over (in attenuated form) to the case when there is a contractual relationship but one based on imperfect knowledge.

²³ We could allow for injurers to choose not just whether to engage in their activity but also at what level to engage in it. Although this would not change the qualitative nature of the results, it would add to the complexity of the analysis.

in the activity. Victims, however, make no decisions; they are passive. Accidents are therefore unilateral in nature. This assumption is descriptive of situations in which whatever changes in the behavior of victims that could reasonably be expected to follow from changes in liability rules would have only a small influence on accident losses. The assumption is also pedagogically convenient, in that the issues concerning causation and the scope of liability are in the main logically tangential to those raised by the possibility of contributory negligence (see Part VII-D).

Parties understand and take as given both the way that the scope of liability is determined and the liability rule—strict liability or negligence—employed by the courts, and the courts are assumed to have perfect knowledge about each accident. Administrative costs associated with the use or threatened use²⁴ of the legal system are borne only in accidents that fall within the scope of liability. The sole importance of this assumption is that administrative costs decline as the scope of liability is reduced.²⁵

Under all these assumptions, comparisons can be made among the resulting outcomes given various ways of determining the scope of liability and given the choice of liability rule. In making comparisons, the following additional assumptions will be used. First, the utility or disutility of any action or loss has a well-defined monetary equivalent. Second, parties are “risk neutral”:²⁶ their *ex ante* evaluation of their position—their “expected utility”—is its expected monetary (plus monetary equivalent) value, and they choose among actions on the basis of this expected value. Third, the measure of social welfare is the

²⁴ Reaching an out-of-court settlement—the way that most disputes are resolved—certainly absorbs resources: the injurer and victim must bargain, perhaps hire lawyers, and possibly begin proceedings as well.

²⁵ If an accident occurs and is not within the scope of liability, one would expect it less likely to be seriously disputed than if it were within the scope of liability. This is the principal way in which it is imagined that administrative costs fall as the scope of liability is restricted. However, it is admitted that, for cases which are seriously disputed, administrative costs could in principle rise as the scope of liability is restricted, since a greater number of escapes for the defendant would have to be considered. If this possibility is felt to be dominant in importance in some context (an unlikely possibility in my opinion) so that administrative costs rise as the scope of liability is restricted, then in our analysis the only reason for restricting the scope of liability would be to avoid problems of crushing liability.

²⁶ This assumption is briefly discussed in Part VII-G *infra*. Let us note here that because the assumption means that parties would not wish to purchase insurance even at actuarially fair rates, we do not have to study the effect of the purchase of liability coverage on the behavior of injurers. However, were we to study this effect, the qualitative nature of our results would not be much altered, and for the following reasons. On the one hand, liability premiums for parties engaged in a particular activity must reflect expected liability costs, so injurers would still be forced to consider such costs in deciding whether to engage in the activity. On the other hand, while an insured injurer's incentive to take care might be diminished, it would generally not be absent, both because his coverage might be incomplete and because of linkage of the premium, claims paid, or renewal of policy to his behavior. (See Steven Shavell, *Accidents, Liability, and Insurance*, *Am. Econ. Rev.* (forthcoming), for an analysis of tort law and insurance.)

sum of expected utilities; equivalently, and as stated previously, it is the sum of benefits parties derive from engaging in their activities minus the costs of exercising care minus expected accident losses and administrative costs.²⁷

IV. THE FIRST-BEST SOLUTION TO THE PROBLEM OF ACCIDENTS²⁸

As explained in the introduction, the first-best solution to an accident problem is what an (all-powerful) authority who wanted to maximize social welfare would command to be done; and it is also what injurers and victims would have agreed to do had they come together and reached a mutually satisfactory bargain at the outset.²⁹

In order to achieve a first-best solution, the authority would have to decide whether injurers ought to engage in their activity and, if so, what their level of care should be. Were injurers to engage in their activity, social welfare would equal the benefit to them from engaging less the cost of taking care less expected accident losses given the level of care. And the level of care that the injurers would be told to take would be the first-best level—the level that minimizes the sum of the cost of taking care plus expected accident losses.

If injurers were not to engage in their activity, social welfare would be determined by whatever expected accident losses would then be, for it is assumed that victims might be involved in accidents in the absence of injurers. If social welfare in this case exceeds social welfare in the former, the authority would not order injurers to engage. Otherwise, it would order the injurers to engage and to take the first-best level of care.

Let us now consider the first-best solution in three examples. The examples and variants of them will be referred to later in Parts V and VI.

Example 3. If a bicyclist rides in a park, he may be involved in an accident with a jogger, who might be involved in an accident in any event. The situation is illustrated in Table 3. If visibility is poor, then, no matter what the bicyclist does, the jogger will be in an accident and bear a loss of 200. Thus, if visibility is poor, the bicyclist's presence in the park will not be a cause in fact of the accident loss of 200. (Let us imagine that if there is no bicyclist about the jogger will trip on a fallen branch and sprain his ankle. If there is a bicyclist riding in the park, he will run into the jogger when he is about to trip on the

²⁷ This figure is equivalent to the sum of expected utilities because parties' expected utilities are comprised of the benefit (if any) from engaging in activities minus the costs of exercising care minus expected accident losses and administrative costs.

²⁸ See part A of the Appendix for a formal presentation of this section.

²⁹ Were victims and injurers to come together at the outset, they would by mutual consent wish to have the sum of expected values (which is social welfare) at a maximum, for otherwise everyone could be made better off by raising the sum and by the gainers then compensating the losers in the altered situation. (In the language of welfare economics, any Pareto-efficient first-best solution is such that the sum of expected values is maximized.)

TABLE 3

ACTIONS	STATES		
	Visibility Poor	Visibility Moderate	Visibility Good
<i>Probability</i>	.01	.02	.97
Bicyclist rides in park but not carefully	loss of 200	loss of 100	no loss
Bicyclist rides in park carefully	loss of 200	no loss	no loss
Bicyclist stays away from park	loss of 200	no loss	no loss

Notes: Cost of riding carefully = 1

Gross benefit to bicyclist of being able to ride in park = $2\frac{1}{2}$.

branch, and the jogger will sprain his ankle just the same.) However, if visibility is moderate and the bicyclist rides carefully, he will manage to avoid an accident with the jogger.

Social welfare may be computed for each of the three actions of the bicyclist.³⁰

Bicyclist rides in park but not carefully:

$$2\frac{1}{2} - (.01)(200) - .02(100) = -1\frac{1}{2}$$

Bicyclist rides in park carefully: $2\frac{1}{2} - 1 - .01(200) = -\frac{1}{2}$.

Bicyclist stays away from park: $-.01(200) = -2$

Thus, social welfare is maximized if the bicyclist rides in the park and takes care.

Example 4. If a factory operates, it releases pollutants that increase the risk of damage to nearby residents (homes might have to be repainted more frequently). There is already a risk due to pollution from other (and, let us suppose, natural) sources. (See Table 4.)

No matter what the level of pollution from other sources, control of emissions will reduce losses due to pollution. If pollution from other sources is moderate or heavy, there will be some losses even if the factory does not operate.

³⁰ The reader should not be disturbed that social welfare is a negative number here (and in some of the other examples as well). This is because expected accident costs happen to be larger than the injurer's benefit, if any, from engaging in his activity. Nor should the reader be disturbed that we have ignored any benefit the victim gets from his activity. Taking account of such a benefit could not alter the results in the examples (by assumption, the victim is passive and cannot alter his behavior); it would just add a constant to social welfare (see also note 80 *infra*).

TABLE 4

ACTIONS	STATES		
	No Pollution from Other Sources	Moderate Pollution from Other Sources	Heavy Pollution from Other Sources
<i>Probability</i>	.25	.5	.25
Factory operates and does not control pollutant emissions	loss of 10	loss of 20	loss of 30
Factory operates and controls emissions	no loss	loss of 15	loss of 20
Factory does not operate	no loss	loss of 10	loss of 15

Notes: Cost of control of emissions = 7

Benefit to society if factory operates = 15. (This benefit equals the gross benefits to consumers of the use of the factory's output less the production cost.)

Social welfare may be computed for the three possible actions of the factory:

Factory operates and does not control emissions:

$$15 - .25(10) - .5(20) - .25(30) = -5$$

Factory operates and controls emissions:

$$15 - 7 - .5(15) - .25(20) = -4.5$$

Factory does not operate: $-.5(10) - .25(15) = -8.75$

Consequently, social welfare is maximized if the factory operates and controls emissions.

Example 5. If a car drives down a country road, it may be involved in an accident with cattle, which occasionally cross the road at either point A or point B. Point B is further down the road than A, and both points are obscured from view. The car may also be involved in an accident with a slow-moving tractor, which sometimes travels down the road. (See Table 5.)

Speeding is a cause of expected losses given that a tractor is present. However, speeding is not a cause of expected losses given that cattle cross the road at *some* point. The speed of the car given that cattle cross the road affects only the position of the car when an accident occurs, a circumstance that is irrelevant to losses. If the car speeds and the cattle cross at B, speeding will be a cause in fact of an accident; and if the car does not speed and the cattle cross at A, not speeding will be a cause in fact of an accident.

Social welfare for the three possible actions of the driver is as follows.

$$\text{Car speeds down the road: } 5 - .02(100) - .03(200) = -3$$

$$\text{Car does not speed down the road: } 5 - 1 - .02(100) = 2$$

$$\text{Car does not go down the road: } 0$$

TABLE 5

ACTIONS	STATES			
	No Cattle or Tractor Present	Cattle Cross at A	Cattle Cross at B	Tractor Present
<i>Probability</i>	.93	.02	.02	.03
Car speeds down the road	no loss	no loss	loss of 100	loss of 200
Car does not speed down the road	no loss	loss of 100	no loss	no loss
Car does not go down the road	no loss	no loss	no loss	no loss

Notes: Cost to driver of not speeding = 1
Gross value to driver of going down the country road = 5

Therefore, social welfare is maximized if the car goes down the country road but does not speed.

It should be clear from these examples³¹ that the general description given above of the first-best solution may be paraphrased by two statements. (a) It is socially desirable for the injurer to engage in his activity if the benefits less the cost of care less the *increment* in expected accident losses over what they would be were he not to engage is greater than zero. (b) The first-best level of care is determined by the cost of taking care and the degree to which lack of care is a cause of expected losses.

V. THE SCOPE OF LIABILITY UNDER STRICT LIABILITY³²

It will be assumed here if an injurer chooses to engage in his activity and an accident occurs that is within the scope of liability, then the injurer will be

³¹ In Examples 3-5 it was always best for the injurer to engage in the activity and to take care. We constructed the examples in that way only to be able to illustrate certain points in later parts of the paper; there is obviously no presumption that in fact it would be best to engage in an activity or to exercise some named level of care.

³² The claims of this part are precisely stated and proved in part B of the Appendix. The reader should therefore not believe that our reasoning here overlooks significant factors (given our assumptions) or that the examples are artfully chosen to illustrate what is in fact special or at least plausibly contradicted. Not only are the claims valid in a general setting, but also their true nature and their very statement is for the most part simpler than what one might conclude from our discussion. This simplicity is achieved by viewing care in the formal analysis as a continuous variable and the probability measure over states of the world also as continuous.

strictly liable—he must pay in damages the victim's losses (of course, if the injurer does not engage in the activity, he will not be liable).³³ Whether an accident is within the scope of liability will be decided by the simple criterion of whether the state of the world that occurred is in a designated set \mathcal{S} (for scope),³⁴ and it will be assumed that this set is selected by the courts to maximize social welfare.

Let us first consider the situation when there are no administrative costs associated with the use or threatened use of the legal system.

In order to understand how the set \mathcal{S} determining the scope of liability should be selected, we must consider the effect of strict liability on the injurer's decision whether to engage in an activity and, if so, on his decision about care.

With regard to the decision whether to engage in an activity let us restate the argument sketched in the introduction about the need to avoid crushing liability. Suppose that the scope of liability is complete: if the injurer engages in his activity, he will be held liable for all accidents. Then he might decide not to engage in his activity when, according to the first-best solution, he should engage in it. He might decide not to engage in it because what he would see as the cost of doing so is the cost of care plus the *total* of expected accident losses. Yet the true social cost of engaging in it is the cost of care plus only the *increment* in expected accident losses over what such losses would be in his absence. Consequently, in order to induce the injurer to engage in his activity when it is socially desirable for him to do so, it may be necessary to restrict the scope of liability, which is to say, restrict the set \mathcal{S} .³⁵

However, the injurer would generally choose a level of care lower than the first-best level if the scope of liability were not complete, for he would not be led to balance the cost of care against the expected reduction in accident losses *no matter how they occurred*, but *only* against the expected reduction in losses for those accidents within the scope of liability. Yet, we should mention an

³³ Not only is a study of causation and the scope of liability under strict liability of interest in its own right, it is also essential to an understanding of causation under the negligence rule, the problem to which most writers on causation have restricted their attention.

³⁴ The interpretation is that the court may consider the particular circumstances of the accident (formally embodied in $\ell(\cdot, t)$ and $\ell(t)$)—see the Appendix). It will be clear from what follows that this implicitly allows for the use of tests corresponding to causation in fact, substantial factor, etc.

³⁵ An alternative would be to restrict not the scope but the extent of liability—by having the injurer pay only the increment in the losses due to the injurer's activity. This possibility is ruled out from our model by assumption. The excuse is twofold. First, as this alternative would require the broadest scope of liability, administrative costs would bulk large. Second, as a practical matter, the determination of exactly what would constitute the correct increment in losses due to the injurer might often be difficult or impossible (see our comments about this in Part VII-H). In any event, the alternative of restricting the extent rather than the scope of liability has not in fact been adopted as a *general* solution to the particular problem at hand.

important qualification to be amplified below. If care would not reduce accident losses given a particular state, then there would be no reduction in care were the state to be excluded from \mathcal{S} . This is exactly the situation when failure to alter the level of care is not a cause in fact of the accident.

To summarize, in order to induce the injurer to engage in his activity when it would be socially desirable that he do so, the scope of liability may have to be restricted, but this will often result in a socially undesirable reduction in the level of care. Let us illustrate these ideas.

Example 6. Refer to Example 3 concerning the bicyclist and the jogger. We want to verify (among other things) two facts. If the scope of liability is complete, then there would be crushing liability and the bicyclist would decide against riding in the park, a socially undesirable outcome. But, if the bicyclist would be held liable only when visibility is moderate—which, recall, is the circumstance where *his not taking care would be a cause in fact of accident losses*—then he would decide to ride in the park and to take care. (Therefore, the first-best solution can be achieved.) Restricting the scope of liability in this way reduces the expected burden of liability sufficiently to induce him to ride in the park but results in no loss of incentive to take care; and the latter is true because the circumstances in which there is no liability are also those in which care has no effect on accident losses.

To demonstrate all this, we must enumerate the different possibilities for the set \mathcal{S} defining the scope of liability, and for each possibility we must consider how the bicyclist would be induced to behave:

\mathcal{S} equals all states; if he rides in the park, the bicyclist is always liable. If he rides in the park but not carefully, his expected utility is $2\frac{1}{2} - .01(200) - .02(100) = -1\frac{1}{2}$. If he rides in the park and exercises care, his expected utility is $2\frac{1}{2} - 1 - .01(200) = -\frac{1}{2}$. If he does not ride in the park, he gets no benefits but is not liable, so his utility is 0. Thus he would decide not to ride in the park.

\mathcal{S} excludes the first state; the bicyclist is not liable when visibility is poor. If he rides in the park but not carefully, his expected utility is $2\frac{1}{2} - .02(100) = \frac{1}{2}$. If he rides in the park and exercises care, his expected utility is $2\frac{1}{2} - 1 = 1\frac{1}{2}$. If he does not ride in the park, his expected utility is 0. Thus he would decide to ride and to exercise care. As was shown in Example 3, this outcome maximizes social welfare.

\mathcal{S} excludes the second state; the bicyclist is not liable when visibility is moderate. If he rides in the park but does not exercise care, his expected utility is $2\frac{1}{2} - .01(200) = \frac{1}{2}$. If he rides in the park and exercises care, his expected utility is $2\frac{1}{2} - 1 - .01(200) = -\frac{1}{2}$. If he does not ride, it is 0. Thus he would decide to ride but not exercise care. (Under this \mathcal{S} , the expected liability cost is reduced by enough to make it worthwhile for the bicyclist to ride in the park; but because when visibility is moderate taking care has an effect on accident

losses, excluding this state results in a diminished incentive to take care.)

We will not consider here the other possibilities for the set \mathcal{S} ; they do not result in desirable outcomes.

Thus, in this example it is indeed best to exclude from the scope of liability accidents occurring when visibility is poor.

Example 7. Referring to Example 4 concerning the polluting factory, we wish to show two facts. First, if the scope of liability is complete, there will be crushing liability and the factory will be driven out of business when it would have been socially desirable for it to operate. And, second, by appropriately restricting the scope of liability, the factory can be induced to operate, but it cannot at the same time be induced to control emissions of pollutants. (Thus, in contrast to the previous example, the optimal achievable solution is inferior to the (hypothetical) first-best solution.)

Before demonstrating this, several additional assumptions about the nature of the market are necessary. Let 20 be the value of the factory's output to consumers and 5 be the production cost (exclusive of the cost of care), so that $20 - 5 = 15$ is indeed the direct benefit to society from operation of the factory. Assume that competitive forces keep price equal to total cost—production cost plus the cost of care plus expected liability costs—and note that price, and thus total costs, cannot exceed 20, the value of output to consumers. If total costs exceed 20, the factory will go out of business. Last, assume that the factory seeks to maximize profits.

Now consider the outcome under different choices of \mathcal{S} :

\mathcal{S} equals all states; if the factory decides to operate, it is always liable. If the factory operates and does not control emissions, its total cost is $5 + .25(10) + .5(20) + .25(30) = 25$. If the factory operates and controls emissions, its total cost is $5 + 7 + .5(15) + .25(20) = 24\frac{1}{2}$. Since in each case costs exceed 20, the factory would not operate. Under this \mathcal{S} , the factory pays for all pollution costs, not just the incremental costs associated with its operation.

\mathcal{S} excludes the first state; the factory is not liable when there is no pollution from other sources. If the factory operates and does not control emissions, its total cost is $5 + .5(20) + .25(30) = 22\frac{1}{2}$. And if it does control emissions, its total cost is $5 + 7 + .5(15) + .25(20) = 24\frac{1}{2}$. Again, the factory would decide not to operate. Under this \mathcal{S} , the restriction of the scope of liability does not lower costs enough to allow the factory to sell at a price at which consumers would make purchases.

\mathcal{S} excludes the second state; the factory is not liable when there is moderate pollution from other sources. If the factory operates and does not control emissions, costs are 15 (we omit calculations), and if it does control emissions, costs are 17. The factory would thus choose to operate but not control emissions. Under this \mathcal{S} , the restriction of the scope of liability lowers costs

enough to allow the factory to sell at a price at which consumers will buy. At the same time the incentive to control emissions is diminished, since there is no inducement to consider the reduction of 5 in accident costs that would result from emission control were pollution from other sources moderate.

\mathcal{S} excludes the third state; the factory is not liable when there is heavy pollution. Under this \mathcal{S} , the factory would operate but not control emissions for reasons similar to those mentioned in the last paragraph.

The other possibilities for \mathcal{S} are worse or no better and will not be considered.

Note from Example 4 that social welfare is higher when the factory operates and fails to control emissions than when it does not operate at all. It is therefore best to restrict the scope of liability (by excluding the second or third state from \mathcal{S}), but this means that the factory is not induced to control emissions.

Example 8. Refer to Example 5 concerning the car traveling down the country road. It may easily be verified (along the lines used in the last two examples) that, if the driver is always held liable, he would decide to drive down the road and not to speed, a decision that is the first-best outcome. That this should be true is obvious. The only reason (at this stage in the analysis) for restricting the scope of liability is to avoid making the injurer pay for accident costs that would occur in his absence, and in this example there are no such costs.

However, if the driver were not held liable for striking down cattle, then he would still decide to drive down the road and not to speed. Restricting the scope of liability in this way does not reduce the incentive to drive slowly because the speed of the car has no influence on the expected accident losses *given* the event that cattle cross the road. The incentive to drive slowly can come about only by making the driver liable for collisions with the tractor. This example is a classic type in which causation in fact (the high speed of the car would be a cause in fact of the accident if the cattle crossed at B, and the failure to go faster if the cattle crossed at A) need not imply liability. Yet, the impression should not be created (as it sometimes is) that there would never be a reason to hold the driver liable for striking down cattle. To understand why, suppose the loss for striking cattle is not 100 but 1000. Then, because expected losses due to accidents involving cattle are so high, the first-best outcome is for the driver not to go down the country road. This outcome would be achieved if the driver is always liable for striking down cattle. In order to discourage the driver from going down the road, the scope of liability must be broadened; the fact of his speed not affecting the possibility of accidents with cattle is simply irrelevant.

Let us now consider the situation when there are positive administrative

costs associated with the use or threatened use of the legal system. As suggested in the introduction and as will be illustrated in two examples, positive administrative costs provide another reason for reducing the scope of liability, namely, a savings in administrative costs.

Example 9. Consider again the case of the car going down the country road. As noted in Example 8, holding the driver liable for all accidents or allowing him to escape liability for striking down cattle were equally desirable alternatives. Either way of defining the scope of liability would result in a first-best outcome. When, however, administrative costs are introduced, we should expect, and are about to verify, that the narrower scope of liability is socially preferable.

Recall from our Part III that we assume administrative costs are borne if and only if an accident is within the scope of liability. In this example, suppose that administrative costs are 10 for the injurer and also 10 for the victim.

Now consider the outcome under various choices for \mathcal{S} :

\mathcal{S} equals all states. Under this \mathcal{S} , the driver would decide to go down the road and not to speed. His expected utility would be $5 - 1 - .02(110) = 1.8$ and social welfare would be $5 - 1 - .02(120) = 1.6$.

\mathcal{S} excludes states 2 and 3—no liability for accidents involving cattle. Here, again, the driver would decide to go down the road and not to speed. His expected utility would be $5 - 1 = 4$ and social welfare would be $5 - 1 - .02(100) = 2$. (The savings in administrative costs is $.02(20) = .4$, for $.02$ is the probability of striking down cattle and 20 the total administrative cost that would have been borne had such accidents been in \mathcal{S} .)

It may also be verified that other possibilities for \mathcal{S} could not be better and we will not consider them. Thus, it is indeed strictly preferable to exclude from liability accidents involving cattle.

Example 10. Table 6 is a modification of Table 3. This example has been designed to isolate the effect of administrative costs on the choice of the scope of liability. Because there are no accidents if the bicyclist does not ride in the park, the issue of crushing liability, which was stressed before, does not enter.

We want to illustrate that liability ought to be restricted to the type of accident situation in which the bicyclist can exercise the greatest effect on accident losses by taking care. In this example, therefore, we want to show that there should be liability only when visibility is moderate (losses would fall by 100 if care were exercised) rather than only when it is poor (losses would fall by 50).

The reader can easily verify the following:

\mathcal{S} equals all states. If the bicyclist is always held liable when he rides in the park, he would decide to ride in the park and to exercise care. Social welfare would be $25 - 7 - .1(50 + 5 + 5) = 12$.

TABLE 6

ACTIONS	STATES		
	Visibility Poor	Visibility Moderate	Visibility Good
<i>Probability</i>	.1	.1	.8
Bicyclist rides in park but not carefully	100	100	no loss
Bicyclist rides in park carefully	50	no loss	no loss
Bicyclist does not ride in the park	no loss	no loss	no loss

Notes: Cost of riding carefully = 7

Benefit to bicyclist of riding in park = 25

Administrative cost borne (by both the jogger and bicyclist) when an accident is in the scope of liability = 5

\mathcal{S} excludes the second state. If the bicyclist is held liable only when visibility is poor, he would decide to ride but not to take care (the reduction of 50 in accident costs is not a sufficient motive). Social welfare would be $25 - .1(100 + 5 + 5) - .1(100) = 4$.

\mathcal{S} excludes the first state. If the bicyclist is held liable only when visibility is moderate, he would decide to ride and to take care (the reduction of 100 in accident costs is a sufficient motive). Social welfare would be $25 - 7 - .1(50) = 13$. This is higher than social welfare when the bicyclist is always liable, owing to the savings in administrative costs.

\mathcal{S} excludes all states. If there is no liability, the bicyclist would decide to ride but not carefully. Social welfare would be $25 - .1(100) - .1(100) = 5$.

Thus, it is best to restrict liability to accidents occurring when visibility is moderate.

The logic of the argument to this point and, especially, the examples suggest much more than the desirability of restricting the scope of liability in some general and unspecified way. They indicate also *how the scope of liability ought to be restricted: namely, to types of accidents with the "fundamental" characteristic that, given the circumstances under which the type of accident occurs, the effect of an increase in care in reducing accident losses should be sufficiently pronounced.* Why this characteristic ought to determine the scope of liability is that liability should be restricted in such a way that the disadvantage of so doing—the reduction in incentives to take care—is felt as little as possible. And to do this, in turn, it is best to eliminate from the scope of liability accidents occurring in circumstances under which taking care would have had a sufficiently small effect on losses.³⁶

³⁶ What constitutes "sufficiently" depends on, among other things, the magnitude of administrative costs. A higher level of administrative costs would imply a smaller scope of liability, and

The fundamental characteristic implies, and our examples illustrate, that *for an accident to be in the scope of liability, the injurer's not having altered his level of care should be a cause in fact of the accident*. We say "implies" because when there is no causation in fact, the effect of an increase in care on accident losses is, by definition, zero, so that it is certainly not "sufficiently pronounced."

The fundamental characteristic also implies that "*coincidental*" accidents *ought not to be included in the scope of liability*. Coincidental accidents are exemplified by speeding cases³⁷ like *Berry* and our Example 5 about a driver who might strike down cattle that cross a country road.³⁸ In these cases, speed is a cause in fact of an accident because, but for the speed, the vehicle would not have been in the precise position necessary for the occurrence of the accident (at the point where the tree happened to be blown over in *Berry*; at the point where the cattle happened to cross in Example 5). But speed is irrelevant to losses on average in accidents of the type in question (speed did not affect expected accident losses involving *some* tree being blown over and then falling on the streetcar; or cattle crossing at *some* point in the road). In other words, in such cases the effect of an increase in care in reducing accident losses is nil given the type of accident. Additionally, and as will be discussed later (Part VII-B), a similar line of reasoning shows that *accidents outside the normal risk created by an act should not be in the scope of liability*.

We have not yet stated what the connection is between the determination of the scope of liability under our theory and the use by the courts of the concept of unforeseeability. As will be argued in Part VII-A, when the theory is appropriately extended and interpreted, it does furnish an explanation of the role of unforeseeability. Nevertheless, what we want to discuss now—because it will provide insight and help to resolve a puzzle—is that the theory as presently developed indicates that, other things equal, the probability of an accident should *not* influence whether it is included in the scope of liability. Consider the following example.

Example 11. For simplicity, let us modify Example 4 concerning the polluting factory and imagine that weather conditions—temperature, winds, and so forth—affect losses given the action of the factory. (See Table 7.)

therefore a more stringent requirement on the effect of care on accident losses before inclusion in the scope of liability.

³⁷ See Hart & Honoré, *supra* note 17, at 114.

³⁸ Let us, however, repeat the qualification noted in Example 8: Apart from whatever care he takes, the injurer, solely by virtue of engaging in his activity, may create sufficiently high expected accident losses that he ought not to engage in it (should not drive down the country road owing to the number of cattle he would be expected to strike down, should not operate a streetcar on a route where there are many trees which are for some reason prone to fall down). In such a situation, therefore, it would be desirable to impose (crushing) liability.

TABLE 7

ACTIONS	STATES		
	Normal Weather Conditions	Bad Weather Conditions	Freakish Weather Conditions
<i>Probability</i>	.9	.09	.01
Factory operates and does not control pollutant emissions	loss of 100	loss of 200	loss of 200
Factory operates and controls emissions	loss of 100	loss of 100	loss of 100
Factory does not operate	loss of 90	loss of 90	loss of 90

Notice that control of emissions reduces losses only when weather conditions are bad or freakish. Notice also that the levels of loss as a function of the firm's actions are the same whether the weather is bad or freakish; the sole difference, therefore, between pollution losses suffered in these two states is that the latter state is relatively unlikely.

Assume that the value to consumers of the factory's output is 50 and that the cost of production is 5. Hence, for the factory to operate, its expenditure on control of emissions plus expected liability payments must not exceed 45. Last, assume that the cost of controlling emissions is 9.1. Under these assumptions the first-best solution is for the firm to operate and to control emissions. This can be achieved if and only if the injurer is held liable for pollution losses occurring in bad *and* in freakish weather conditions but not in normal conditions. Moreover, the example could have been altered so that the probabilities of the last two states were .099 and .001 or .0999 and .0001, etc., without changing the result that losses occurring in freakish conditions should be in the scope of liability.

The intuition behind the result just illustrated is straightforward. While it is true that the exclusion of a low probability accident from the scope of liability would result in a small loss in incentives to take care, it is also true that it would result in a small decrease in the injurer's expected liability payments and also a small decrease in expected administrative costs. These two effects are exactly offsetting, meaning that low probability does not work against (or in favor of) inclusion in the scope of liability.

Indeed, the conclusion could not have been otherwise. As others have stated, the assumption that low probability accidents ought to fall outside the scope of liability leads to a *reductio ad absurdum*. If we describe any accident in sufficient detail, we will have to conclude that its probability was extremely small. (The likelihood was negligible that *this* particular injurer

struck down *that* particular victim on Tuesday at 3:48 P.M. at *that* location when the temperature was 44°F) Thus, the assumption would lead us to say that there should not be liability for any accident whatever.³⁹

Summary of the Argument

1. The advantages of restricting the scope of liability are as follows. (a) It reduces the burden on injurers, thereby avoiding crushing liability: if made to pay for most or all accidents in which they are involved, injurers would be forced to bear accident losses that would have occurred in their absence. Therefore, they might find their activity too costly to engage in even though it might be socially desirable for them to do so. (b) It reduces administrative costs associated with the use or threatened use of the courts.

2. The disadvantage of restricting the scope of liability is the reduced incentive to prevent accidents: if injurers have to pay for fewer accidents, they will take less care to prevent them (but see the important qualification to this in 4(a) below).

3. The social-welfare-maximizing scope of liability appropriately balances the advantages and the disadvantage. Their relative importance determines the scope of liability in a general, overall sense.

4. *For a type of accident to be included in the scope of liability, it should have the fundamental characteristic that, given the circumstances under which it occurs, there must be a sufficiently high potential for reducing losses by taking more care.* If the effect of care on losses is too low, then by excluding the type of accident from the scope of liability there is little reduction in the incentive to take care but there are the two advantages of having restricted liability. The fundamental characteristic implies the following.

(a) Causation in fact is a prerequisite for inclusion in the scope of liability. If failure to alter the level of care is not a cause in fact, then, by definition, there is no potential for reducing losses by taking more care.

(b) Coincidental accidents and those falling outside the normal risk created by an act should not be in the scope of liability. In coincidental accidents (exemplified by speeding cases like *Berry*) the injurer's failure to take care may be a cause in fact, yet care does not affect expected losses for accidents of that type; and similar reasoning applies to accidents that are not within the normal risk (exemplified by the case of the gun that fell on the child's toe).

(c) The probability of an accident should have no bearing on inclusion in the scope of liability. The probability of an accident has nothing to do with whether it is described by the fundamental characteristic. Or, to put it

³⁹ Compare the discussions (and references cited therein) of William L. Prosser, *supra* note 3, at 267-68; and Hart & Honoré, *supra* note 17, at 232.

another way, while it is true that if a low probability accident is excluded the reduction in the incentive to take care would be small, so would be the two advantages of restricting the scope of liability. Or, to put the matter still another way, a conclusion contrary to ours leads to a well-known *reductio ad absurdum*, for any accident when described in sufficient detail is seen to be extremely unlikely. This puzzle, our result, and the fact that many unlikely accidents are in fact excluded from the scope of liability will all be resolved when unforeseeability is discussed in Part VII.

VI. THE SCOPE OF LIABILITY UNDER THE NEGLIGENCE RULE⁴⁰

It will be assumed in this part that, if an injurer decides to engage in his activity and an accident occurs which is within the scope of liability, then the injurer will be required to pay damages if he was negligent. That is, the injurer will pay damages if and only if two conditions are met: the state of the world that occurred is in the set \mathcal{S} defining the scope of liability; and the level of care exercised by the injurer falls short of the "due care" level.⁴¹ It will be assumed that \mathcal{S} and the due care level are chosen by the courts to maximize social welfare. And since we assume the courts have perfect information about accidents, they make no "mistakes." They correctly calculate the level of due care; they "observe" with complete accuracy the level of care actually exercised by injurers, and so on. The importance of this assumption to our conclusions in this part is great, as will be explained subsequently.

As in the previous part, we first examine the situation in the absence of

⁴⁰ See Part C of the Appendix for a formal presentation of this part.

⁴¹ Although the question whether a meaningful distinction can be drawn between the determination of fault under the negligence rule and the determination of the scope of liability (presumably by reference to causal principles) is sometimes raised, in this model there is no ambiguity between the two; it is simply logically incorrect, *wrong*, to think otherwise. Fault is determined in the model by whether the level of care falls short of due care. The scope of liability is determined by the circumstances of the accident (by whether the state of the world is in \mathcal{S}). There is no necessary connection between the two. For example, fault could be determined by the Hand formula and the scope of liability by the calendar date.

However, that the question is occasionally raised is understandable. On the one hand, in attempting to find the appropriate due care level, the focus is naturally on the likely and important effects of the exercise of care. It is relatively unimportant to the determination of the correct due care level how care would alter accident losses in very unlikely circumstances. Moreover, there is no effect on the appropriate due care level attributable to the possibility of accidents where the exercise of greater care would have no effect (accidents where there is no but-for causation, accidents lying outside the normal risk, etc.). On the other hand, accidents occurring in these latter types of circumstance are those which are in fact, and in our model ought to be, outside the scope of liability. In other words, the types of accidents which do not raise the rationally determined due care level are frequently those lying outside the rationally determined scope of liability. This is quite a different statement from the assertion that as a matter of logic the fault determination cannot be distinguished from the causal one.

administrative costs associated with the use of the legal system. To determine the scope of liability we must consider how the negligence rule influences the injurer's decision whether to engage in an activity and, if so, how it influences his choice about the level of care.

In deciding whether to engage in his activity, the injurer will compare the value of not engaging to that of engaging and exercising due care. (Why we assume that he would wish to exercise due care will be explained below.) If he engages and exercises due care, he will not have to pay for accident losses that might occur (the court always discovers the truth). This means that he will not be motivated to take into account the increment in expected accident losses occasioned by his engaging in the activity. Consequently, we might expect him to engage in the activity even when it would be socially desirable that he does not—in complete contrast to the fact that, under strict liability, he might not engage when it would be desirable that he do so.⁴² Restricting the scope of liability would not solve this problem; it would only make it worse.

When deciding whether to exercise due care (clearly, he will not wish to exercise more than due care), the injurer will have to consider the consequences of exercising less than due care, of acting in a negligent way. We will assume that the due care level is set low enough that the injurer would decide against acting in a negligent way, for otherwise, we would in effect be considering again a situation in which the injurer is strictly liable. Because restricting the scope of liability increases the relative attractiveness of acting in a negligent way, it lowers the due care standard to which an injurer can be induced to adhere. To see this, it may be helpful to think of the extreme case: if the scope of liability is virtually nonexistent, then injurers could not be motivated to raise care to any due care level exceeding the level of care that they would exercise were there no liability.

In summary, restricting the scope of liability exacerbates the problem that an injurer may engage in his activity when he ought not, and it also reduces his incentive to take due care. Thus, in the absence of administrative costs, there is no reason to restrict the scope of liability.

Example 12. Let us modify the example of the jogger and the bicyclist as in Table 8. If the bicyclist rides in the park, it is socially desirable that he rides carefully, for the cost of care is 10 whereas the expected reduction in accident costs by taking care is $.05(120) + .05(100) = 11$. If he rides and takes care, total social costs are $10 + .05(40) = 12$. Thus, if the benefit to the bicyclist is only 11, under the first-best solution he would not ride in the park; but if the benefit to him is 15, he would.

⁴² This theme is developed in Steven Shavell, *Strict Liability versus Negligence*, 9 J. Legal Stud. 1 (1980).

TABLE 8

ACTIONS	STATES		
	Visibility Poor	Visibility Moderate	Visibility Good
<i>Probability</i>	.05	.05	.9
Bicyclist rides in park but not carefully	loss of 160	loss of 100	no loss
Bicyclist rides in park carefully	loss of 40	no loss	no loss
Bicyclist stays away from the park	no loss	no loss	no loss

Notes: Cost of riding carefully = 10

Gross benefit to bicyclist of being able to ride in the park = either 11 or 15 (both cases will be considered)

The reader may now verify the following, given use of the negligence rule with a due care standard of 10; that is, the bicyclist must have been riding carefully to avoid a finding of negligence. (a) If the scope of liability is not restricted, the bicyclist will decide to ride in the park and to exercise due care whether he gets benefits of 15 or of 11—whether or not it is socially desirable that he ride in the park. (b) If the scope of liability is restricted to accidents occurring in only one of the first two states of the world, then the bicyclist will still always decide to ride in the park but will not be motivated to take due care. It is thus undesirable to restrict the scope of liability.

Savings in administrative costs that can be accomplished by restricting the scope of liability must be weighed against the disadvantage of a reduction in the due care standard to which injurers could be induced to adhere. (There is also a related, if somewhat subtle disadvantage: a reduction in the due care standard accentuates the problem that injurers may have too great an incentive to engage in the activity in the first place.) To illustrate this, let us reconsider the previous example.

Example 13. Example 12 is modified in Table 9. Given this situation, the first-best solution is for the bicyclist to ride in the park and to exercise care. The following may also be verified. (a) If there is no restriction of the scope of liability, then under the negligence rule with a due care level of 2.25 the bicyclist would decide to ride carefully. Social welfare would be $10 - 2.25 - .05(80 + 2) - .05(10 + 2) = 3.05$. (b) If the scope of liability is restricted to accidents occurring when visibility is poor, the bicyclist would still decide to ride carefully. Social welfare would be $10 - 2.25 - .05(80 + 2) - .05(10) = 3.15$. Welfare is higher in (b) than (a) due to a savings of 2 in administrative costs if an accident occurs when visibility is moderate. (c) If the scope of

TABLE 9

ACTIONS	STATES		
	Visibility Poor	Visibility Moderate	Visibility Good
<i>Probability</i>	.05	.05	.9
Bicyclist rides in park but not carefully	loss of 100	loss of 40	no loss
Bicyclist rides in park carefully	loss of 80	loss of 10	no loss
Bicyclist keeps away	no loss	no loss	no loss

Notes: Cost of riding carefully = 2.25

Gross benefit to bicyclist of being able to ride in the park = 10

Administrative cost borne (by both jogger and bicyclist) whenever the accident is in the scope of liability = 1

liability is restricted to accidents occurring when visibility is moderate, the bicyclist would decide to ride but not carefully. Social welfare would be $10 - .05(100) - .05(40 + 2) = 2.9$. (d) If there is no liability, the bicyclist would obviously decide to ride but not carefully. Social welfare would be $10 - .05(100) - .05(40) = 3$.

Consequently, it is best to restrict liability to accidents occurring when visibility is poor. This saves on administrative costs and still induces the bicyclist to take care. Restricting liability to accidents when visibility is moderate also saves on administrative costs but fails to induce the bicyclist to take care. Notice here two apparent differences between the poor-visibility situation and that of moderate visibility. When visibility is poor, the absolute level of accident losses is higher than when visibility is moderate; also, the reduction in accident losses due to care is *lower* (20 rather than 30). As will be explained, precisely because the size of losses is higher with poor visibility it is advantageous to include accidents occurring in this circumstance in the scope of liability. The smaller effect of care when visibility is poor does not make inclusion of such accidents in the scope of liability any less desirable.

This illustrates that when administrative costs are positive, restricting the scope of liability may be socially desirable. It also suggests (and is proved in the Appendix) that the sole criterion for including an accident in the scope of liability should be the size of accident losses. Why this should be so is easily explained. Other things equal, it is best to restrict the scope of liability in such a way that the incentive to take due care is reduced as little as possible. To do this, it is best to make as large as possible the liability that the injurer would bear *if* he were to act negligently; thus, those accidents involving the

largest losses are the ones that ought to be included within the scope of liability. Notice how this line of reasoning resembles and how it departs from that applying under strict liability. Under strict liability, as under the negligence rule, it was desirable to restrict the scope of liability so as to reduce the incentive to take care as little as possible. But, to accomplish this under strict liability, it was best to hold injurers liable precisely when the exercise of care would have an effect on accident losses, for having to *pay* for accident losses induced them to take care.

Let us now try to explain why, contrary to our present result of the magnitude of loss determining the scope of liability, in reality the scope of liability under the negligence rule is determined by the body of principles deduced as desirable in the previous part on strict liability. In reality a significant element of strict liability is inherent in the negligence rule: injurers who do in fact act with due care or attempt to do so will sometimes be found negligent and have to pay damages. Courts might make mistakes in formulating due care standards or in gathering facts; injurers might err in deciding how much care to exercise; and an often probabilistic relationship exists between the care injurers generally endeavor to exercise and their momentary behavior.⁴³ Introducing such possibilities into our model⁴⁴ would have made all the conclusions reached about the scope of liability under strict liability relevant under the negligence rule.

Summary of the Argument

1. The advantage of restricting the scope of liability is a reduction in administrative costs associated with the use or threatened use of the courts.
2. The disadvantage of restricting the scope of liability is a reduction in incentives to take due care: the less often an injurer has to pay damages, the less reason he has to exercise due care.
3. The social-welfare-maximizing scope of liability would balance these two factors and thus include an accident in the scope of liability if and only if magnitude of loss is sufficiently high. But this is clearly unrealistic.
4. However, once one recognizes that under the negligence rule there is an important element of strict liability (for a variety of reasons a party attempting to exercise due care may be found negligent), all the results concerning the scope of liability under strict liability become relevant under the negligence rule. In the end, therefore, the actual role of causation in fact, of coincidence, and so forth in determining the scope of liability under the negligence rule is rationalized but can only be understood by reference to the situation under strict liability.

⁴³ See Peter Diamond, *Single Activity Accidents*, 3 J. Legal Stud. 107 (1974).

⁴⁴ Some of these possibilities are taken into account in *id.*; and in Jerry Green, *On the Optimal Structure of Liability Laws*, 7 Bell J. Econ. 553 (1976).

VII. FURTHER CONSIDERATIONS: THE THEORY EXTENDED

The goal of the previous two parts was to analyze causation and the scope of liability in a context sufficiently simple and idealized to enable us to isolate the factors of greatest conceptual importance. The conclusions drawn will offer general guidance in this part for informally extending and modifying the theory in a way allowing a fuller and more realistic consideration of the issue of concern.

A. *Unforeseeability*

Recall that, according to the theory of Parts V and VI, the probability of an accident should not influence whether it is included within the scope of liability. Moreover, the opposite conclusion would have presented the conundrum that, since any accident is seen to be unlikely when described in fine enough detail, *all* accidents ought to be excluded from the scope of liability.

How can this be reconciled with the fact that in practice unusual, abnormal, freakish—in short, unforeseeable—accidents are often excluded from the scope of liability?⁴⁵ The answer has to do with the possibility that an injurer may underestimate the probability of a type of accident; the subjective probability of the accident which an injurer uses in making his decisions may be less than the subjective probability used by the courts in making their calculations of social welfare.⁴⁶

In order to see what difference an injurer's underestimation makes, consider first the extreme case in which his probability is zero—the possibility of the accident is ignored by him—whereas the courts' probability is positive. In this case, inclusion of the accident in the scope of liability would not have any effect on the injurer's behavior—for his behavior is determined by his probability—yet inclusion would increase expected administrative costs in the courts' calculation—for their probability is positive. Hence, from the courts' point of view, it would be best to exclude the accident from the scope of liability. This conclusion also follows when, instead of a zero probability, the injurer assumes the probability to be positive but less than the courts'. Here inclusion of the accident would have a small effect on the injurer's

⁴⁵ See, for example, the discussions and citations in Prosser, *supra* note 3, at 250-70 and Hart & Honoré, *supra* note 17, at 230-56.

⁴⁶ In Parts IV-VI, we had implicitly assumed that the probability assessments of the injurers and of the courts were identical. In this subpart, the reader may wish to think of the courts' probabilities as "best-informed." However, our argument depends only on the assumed divergence in the two assessments of probabilities (actually, it is their ratio that matters). Also, the argument is not subject to the *reductio ad absurdum* (nor could it be, since it is logically deduced).

behavior relative to the effect it would have if the injurer's probability were as high as the courts', and so forth.

The so-called unforeseeable accidents are just the type for which injurers are likely to underestimate the probability. To argue this, we must consider the habits of thought that guide the injurer's decision in the face of uncertainty. It seems reasonable to suppose that, in making a decision, an injurer does not contemplate each and every one of the vast multitude of possible accidents. Rather, he amalgamates the accidents into a relatively small number of categories and evaluates possible decisions with reference to these categories and their probabilities. In the process, some accidents are completely overlooked since they do not fit into any of the categories (and some categories are assigned an inadequate probability).⁴⁷ On grounds of casual empiricism, it might be said that the accidents for which this is true are often those we call unforeseeable. Psychological studies indicate that individuals tend to underestimate the likelihood of events that are difficult to imagine, or that lie outside the normal realm of experience even if not hard to imagine, or that lack "memorability" even if not hard to imagine and are within the normal realm of experience.⁴⁸

In the celebrated *Palsgraf* case, the issue of unforeseeability was raised in a vivid way.⁴⁹ Several employees of a railroad dislodged a package of fireworks from the arms of a passenger whom they were helping alight an already moving train. The package fell under the wheels of the train and the resulting explosion allegedly caused scales to fall on and to injure Mrs. Palsgraf, who was standing on another platform at the railroad station.⁵⁰ Judge Cardozo, speaking for the majority, found for the defendant in this case. In the terms of our discussion here, his result might be justified by the assumption that the railroad's employees discounted the possibility of the type of accident such as did harm to Mrs. Palsgraf.⁵¹

⁴⁷ For example, the driver of a car might well think of four categories of accident—involving either (1) other moving vehicles or (2) pedestrians or (3) dogs and cats or (4) stationary inanimate objects. Because accidents in which the driver strikes down an escaped animal (leopard, giraffe) from the zoo do not fit in any of the categories, he would ignore their possibility.

⁴⁸ See the well-known article by Amos Tversky & Daniel Kahneman, Judgment under Uncertainty: Heuristics and Biases, 185 *Science* 1124 (1974). See also Daniel Kahneman & Amos Tversky, Prospect Theory: An Analysis of Decision under Risk, 47 *Econometrica* 263 (1979); Paul Slovic, Baruch Fischhoff, & Sarah Lichtenstein, Cognitive Processes and Societal Risk Taking, in *Cognition and Social Behavior* 165 (J. S. Carroll & John W. Payne eds. 1976); and Paul Slovic, The Psychology of Protective Behavior, 10 *J. Safety Research* 58 (1978).

⁴⁹ *Palsgraf v. Long Island R.R.*, 248 N.Y. 339, 162 N.E. 99 (1928).

⁵⁰ We mention parenthetically that it has been argued from the Record of the case that the scale must have been knocked over by a crowd of frightened passengers rather than from the force of the explosion. See William L. Prosser, *Palsgraf Revisited*, 52 *Mich. L. Rev.* 1, 3 (1953).

⁵¹ Cardozo justified his decision on the ground that the actions of the two employees did not constitute negligence *in relation to Mrs. Palsgraf*, and this in turn was true because "if no hazard was apparent to the eye of ordinary vigilance, an act innocent and harmless, at least to

But *Palsgraf* also points up two difficult problems that the courts must face generally in applying a test of foreseeability. First, they must determine whether the injurer really did underestimate the probability of the category of accident. This requires an inquiry into his state of mind. Second, they must (or ought to) recognize that the very use of a foreseeability test reduces the incentive of parties to investigate or, at least, to carefully contemplate the potential consequences of their actions.

B. *The Risk Theory—The “Same Hazard”*

The exclusion of certain accidents from the scope of liability is sometimes explained under the “risk theory” by saying that they were not the result of

outward seeming with reference to her, did not take to itself the quality of a tort The risk reasonably to be perceived defines the duty to be obeyed” 248 N.Y. 342; 162 N.E. 99 (1928). Apparently, Cardozo felt the risk of the type of accident that occurred was not perceived, but there is, as far as I can see, no statement in his opinion to the effect that the misperception of risk is significant *because* it means the incentive to take care would not be affected by a holding of liability in like cases. Thus, the argument of this part is not explicit in Cardozo’s opinion. However, I am inclined to think he would not disagree with it as *an* element to be considered.

In any event, under our general theory (and under most any instrumentalist theory I could imagine) Cardozo’s “relational” notion of negligence seems wrong, and for reasons that are unrelated to misperception of risk. To see this, consider the following example in which accidents might occur involving either a relatively likely victim or a very unlikely victim. Assume that the risks of harm to each type of victim are correctly perceived and that there are just three levels of care which affect the magnitude of loss as shown.

Level of care	Harm if the likely victim is injured—which happens with probability 10%	Harm if the unlikely victim is injured—which occurs with probability 1%
None	1000	1000
Moderate—involving a cost of 5	50	500
High—involving a cost of 10	10	100

What would be due care in relation to the likely victim? The answer to this question based on Cardozo’s notion requires us to think *only* about the likely victim. Thus, due care in relation to him—our duty to him—is the moderate level of care. It is clearly worth the cost of 5 to reduce losses from 1000 to 50 with probability 10% (since $.1(950) = 95$, which exceeds 5). And it is not worth an additional expenditure of 5 to further reduce losses from 50 to 10 with probability 10% (since $.1(40) = 4$, which is less than 5). Similar calculations show that due care in relation to the unlikely victim is also the moderate level. Thus, if the injurer exercised the moderate level of care and, say, the unlikely victim suffered a loss of 500, we should state on our interpretation of Cardozo’s relational theory, “No liability; there was no negligence toward the victim.”

But what is the desirable due care level? It is the *high* level, for increasing care from 5 to 10 reduces losses for *both* types of accident. (The expected incremental benefit is $.1(40) + .01(400) = 8$, which exceeds the incremental cost of 5.) *The desirable level of due care reflects the possibility of all manner of accident; the relational notion ignores this fundamental fact.*

the same hazards or risks which the injurers' actions would normally be thought to create, or they had a coincidental character. Recall the example of a child who was given a loaded gun with which to play and, instead of accidentally shooting a friend (the normal risk created), dropped it on his toe and thereby suffered an injury. In *Berry*, the streetcar was traveling at excessive speed and, instead of colliding with another vehicle or injuring a pedestrian or jumping off the tracks, was hit from above by a falling tree. In the kitchen of a restaurant, an unmarked box of rat poison was placed on a shelf along with foodstuffs and, instead of later being mistakenly used in the preparation of a meal, exploded because, unbeknownst to anyone, it was highly flammable, and the shelf happened to be near a stove.⁵² A holding of no liability in such cases can be explained by the risk theory.

A holding of no liability can also be explained by our theory, in a way consonant with the risk theory; the risk theory is derivative of our theory. Accidents that are not "within the risk" are usually of a type such that the injurer's level of care would have relatively little effect on expected losses. Giving a child a loaded gun rather than an unloaded gun does not raise expected losses from accidents of the type in which the child drops the object with which he is playing on his toe. Driving a streetcar at excessive speed has little or no influence on expected accident losses due to trees falling on streetcars. Placing poison on a shelf along with other foodstuffs does not alter the expected accident losses from explosions of the poison (the probability that the shelf with foodstuffs is near the stove is not greater, or at least not much greater, than the probability that another shelf, say the next, is near to it). Since accidents not within the risk are such that the injurer has relatively little influence on expected accident losses, the theory suggests that they should be excluded from the scope of liability (see especially item 4 in the summaries of Parts V and VI).

C. *Multiple Injurers*⁵³

Let us make two remarks about the situation when more than one injurer is involved in an accident. First, in principle this is virtually always true—even if by "involved" we mean that a party is a cause in fact (much less a probabilistic cause or a cause of expected losses); usually numerous parties could be considered to be involved in some way in an accident. For this reason and because costs of resolving disputes rapidly increase (probably

⁵² This is a hypothetical based on *Larrimore v. American National Insurance Co.*, 184 Okla. 614, 89 P.2d 340 (1939) and given in Robert E. Keeton, *Legal Cause in the Law of Torts* 3 (1963).

⁵³ For an interesting analysis of the functioning of liability rules and of rules regarding contribution when harms are done by several parties, see William M. Landes & Richard A. Posner, *Joint and Multiple Tortfeasors: An Economic Analysis*, 9 J. Legal Stud. 517 (1980).

more than proportionately) with the number of parties, the exclusion of all but one or several of the potential defendants from the scope of liability is usually desirable. When considered in this light, our earlier argument for the rationality of the general requirement of causation in fact is reinforced. Insistence on causation in fact is a simple and effective way of reducing the number of defendants, often without reducing incentives to take care. What alternative requirement would work comparably well is not obvious. Additionally, given causation in fact, it makes sense that the scope of liability should be further restricted to those types of parties who play relatively important and immediate roles in an accident, for their actions are most worthwhile controlling and also are most responsive to imposition of liability.

Our second remark concerns the often-noted cases in which the requirement of causation in fact is relaxed. There are two types of such cases to consider. Suppose it is known that exactly one of several injurers did harm to the victim, but the guilty party is not known.⁵⁴ In *Summers v. Tice*, two hunters using the same type of gun and of birdshot negligently fired at quail when a third hunter, Tice, was nearby. Tice was struck in the eye by a single piece of shot.⁵⁵ Since it was impossible to say who had hit Tice in the eye, neither hunter could be said to be a cause in fact of the injury. The court, however, held both responsible. To do otherwise would inappropriately weaken the incentive of injurers to avoid harm, for to them, but not to society, accidents in which the identity of the injuring party is not known would then be costless.⁵⁶ In the second type of case, the failure of any of

⁵⁴ For relevant citations to cases, see Prosser, *supra* note 3, at 243.

⁵⁵ Another piece of shot hit Tice in the lip but caused no real injury. *Summers v. Tice*, 33 Cal. 2d 80, 199 P.2d 1 (1948).

⁵⁶ The proof is as follows. Assume that the cost of the injury to Tice is purely monetary and amounts to h . Assume also that the two other hunters are identical. Let the probability that a hunter fires and hits Tice be $p(x)$ where x is his level of care and suppose for simplicity that the probability that both hunters hit Tice is zero. Then the first-best problem, to minimize the total cost of care plus accident losses, is to minimize over x

$$x + x + (p(x) + p(x))h = 2x + 2p(x)h \quad (1)$$

Let x^* denote the first-best value of x .

Now suppose that the liability rule is strict liability (analysis of the negligence rule is similar); that a hunter is held liable for h if he hits Tice and the court knows this; that each hunter pays half the damages if it cannot be determined which one hit Tice (in fact, the damages were apportioned); and that λ is the conditional probability that it cannot be determined who hit Tice. Finally, assume that the hunters jointly agree on their level of care. Then their problem is to minimize over x .

$$2x + 2(1 - \lambda)p(x)h + 2\lambda p(x)(h/2 + h/2) = 2x + 2p(x)h, \quad (2)$$

so the first-best outcome is achieved. However, if the hunters would not have to pay, were it impossible to say who hit Tice, their problem would be to minimize over x

$$2x + 2(1 - \lambda)p(x)h \quad (3)$$

so, assuming convexity of p , x would be chosen too low.

several injurers to take care would alone be a sufficient condition for the harm to the victim. Thus none of the injurers is a cause in fact; had any one (but only one) of them taken care, the harm would still have been done. If liability were not found in such cases, the consequence would be to create a motive for injurers to act in concert. My fire spreads and threatens to burn a neighbor's home. I then convince another neighbor to set a second fire that will merge with mine, so that I (and my confederate neighbor) escape liability. Curiously, according to our theory, the possibility of such collusive behavior is the only apparent reason for imposing liability when neither injurer is a cause in fact. Suppose that two individuals are known always to set their fires independently. If their fires happen to spread, merge, and destroy a home and if either fire would have done the same, the theory suggests neither should be liable!⁵⁷ This is for essentially the same reason that an individual ought not be held liable if his fire merges with another fire started by lightning and burns a home.⁵⁸

Although it was assumed here that the hunters choose x together, a similar result holds if they act independently. It would still be true that, when the injurer could not be determined, holding both hunters liable would be superior to exonerating them, but it would no longer be true that the first-best outcome would be achieved were both hunters held liable.

⁵⁷ Of course, the courts have held otherwise. *Anderson v. Minneapolis, St. P. & S.S.M.R. Co.*, 146 Minn. 430, 179 N.W. 45 (1920).

The proof that neither ought to be held liable is as follows. (I wish to thank Robert Rosenthal of Bell Laboratories for correcting my first "proof," establishing that both should be liable.) Assume that the home is worth h ; that the probability of individual i 's fire spreading and burning the home is $p_i(x_i)$, where x_i is i 's level of care; that the event of one fire spreading is independent of the event that the other fire does, so $P_1(x_1)p_2(x_2)$ is the probability that the fires merge and together burn the home. The first-best problem is to minimize the total cost of care plus expected accident losses, to minimize over x_1 and x_2

$$x_1 + x_2 + [p_1(x_1) + p_2(x_2) - p_1(x_1)p_2(x_2)]h. \quad (1)$$

The expression in brackets is the probability that the house burns down. Let x_i^* denote the first-best value of x_i . It is clear that x_i^* is determined by minimizing (1) over x_i , given that $x_j = x_j^*$ ($j \neq i$). Equivalently, x_i^* is determined by minimizing over x_i

$$x_i + [p_i(x_i) - p_i(x_i)p_j(x_j^*)]h. \quad (2)$$

Now suppose that each individual is strictly liable for burning down the house only when his fire does not merge with another one (the analysis under the negligence rule is similar). Then, given that j chooses x_j^* , i 's problem is to minimize (2) over x_i . Thus, x_1^* and x_2^* are chosen in Nash equilibrium when there is no liability for either individual if their fires merge. On the other hand, it is easy to verify that the first-best outcome is not achieved in Nash equilibrium if there is liability when the fires merge. Thus, neither party should be liable when their fires merge and destroy the home.

⁵⁸ In *Cook v. Minneapolis St. P. & S.S.M.R. Co.*, 98 Wis. 624, 74 N.W. 561 (1898), the court held that where the defendant's fire merges with one not attributable to human agency, there is no liability. However, as noted in the discussion of the case in *Hart & Honoré, supra* note 17, at 218, the same court later held that when a fire is of unknown origin, it shall be treated as one of human agency, in which case the defendant is held liable.

D. *Contributory Behavior of Victims*

That victim behavior may contribute to the occurrence of accidents should have no effect on the qualitative nature of our arguments concerning the desirability of restricting the scope of liability. From the point of view of inducing victims to act appropriately, it makes no difference whether a type of accident is included within the scope of liability. If it is included, then, because victims will wish to avoid being found contributorily negligent, they will be induced to avoid the type of accident. And if it is not included within the scope of liability, then, because the victims will wish to avoid bearing their own losses, they will again be induced to avoid the type of accident.

However, suppose that for some reason the desire not to be found contributorily negligent is less likely to induce victims to avoid accidents than is the desire not to bear their own losses. For example, suppose that, for a certain category of product liability accidents, victims know that the courts ignore all but gross instances of contributory negligence. Then there would be an advantage to restricting the scope of liability to types of accidents in which victims could not have played an important contributory role.

A different but related issue concerns the appropriate scope of the defense of contributory negligence.⁵⁹ This is illustrated by a case in which the plaintiff was warned not to stand at the end of a platform because it was slippery with ice.⁶⁰ The plaintiff stood there, nevertheless, and happened to be struck by bricks from the wall above him which was in bad repair due to the defendant's negligence. It was held that the defense of contributory negligence could not be successfully asserted. This holding is rational by an argument analogous to that concerning the risk theory. Were like defendants allowed to claim the defense in like instances, their motive to reduce the chance of accidents (keep the wall in good repair) would be unduly weakened, whereas victims still have sufficient reason to take care if they are prevented from being awarded damages on grounds of contributory negligence only in those instances in which the accident was within the risk created by their action.⁶¹ The appropriate scope of the defense of contribu-

⁵⁹ In this regard, we could have discussed the doctrine of "last clear chance," but it does not fall under the traditional rubric of causation.

⁶⁰ *Smithwick v. Hall & Upson Co.*, 59 Conn. 261 (1890), cited in Hart & Honoré, *supra* note 17, at 192.

⁶¹ The point may be verified using the following model. Let x be the care taken by the defendant to prevent bricks from falling, $p(x)$ the probability that they fall on the employee, h the loss that would result, y the care that the employee takes to avoid slipping on the platform, $g(y)$ the probability that he does so, and k the loss that would then result. Notice that p is a function of x alone; it is independent of y (capturing the notion that the employee's care does not alter the probability of being hit on the head by a brick). Similarly, g is not a function of x . Further, assuming for simplicity that both types of accidents cannot occur simultaneously, the first-best problem is then to minimize over x and y

tory negligence will not be pursued further here, but clearly the issue can be analyzed using our general approach.

E. *Intervening Causes*

Causes that come to bear after the injurer acts but before harm is done are described as "intervening." A complete answer to the question, "What is the relationship between intervening causes and the scope of liability?" must be, "Virtually the whole of this paper." In strict logic an intervening cause is nothing more than the occurrence of some event, any event. However, the term intervening cause connotes the occurrence of an event that is strongly outside the control of the injurer; indeed, the discussion of intervening cause in the literature conveys that impression, for the discussion is not coextensive with the whole subject, causation and the scope of liability.

It seems useful to distinguish intervening causes of natural origin from those attributable to human agency. What is usually meant by the former is the occurrence of some unlikely natural event (see our comments on states of the world as causes at the end of Part II), in which case the question of unforeseeability is at issue. Or else, what is meant is the occurrence of an event outside the risk, in which case the application of the risk theory is at issue.

As to intervening causes attributable to human agency, questions concerning unforeseeability or the risk theory are, again, often at issue. However, frequently at issue also is the additional consideration that it is socially desirable to induce the other human agency to avoid doing harm; thus, liability in many types of cases can be explained by looking at how important it is to give other parties an incentive (see our discussion about multiple injurers). Criminal or intentional acts of parties other than the defendant would seem more important to discourage than those involving uncomplicated negligence, and the former but not the latter tend to exclude the defendant from the scope of liability.⁶² Acts of parties other than the defen-

$$x + p(x)h + y + g(y)k. \quad (1)$$

Denote by x^* and y^* the first-best values of x and y . Now assume that the liability rule is, say, strict liability with a defense of contributory negligence. Let the due care standard be $\bar{y} = y^*$. Assume also that a certain fraction f of the time the court will find the employee contributorily negligent if in fact he was not, and that he would always be found contributorily negligent if he was so. Now suppose that the defense of contributory negligence can only be asserted if the employee sustains harm by slipping on the platform. Then it can be verified that $x = x^*$ and $y = \bar{y} = y^*$ is a Nash equilibrium. However, if the defense may be asserted by the defendant even if harm is sustained by a brick falling on the employee, then (since under our assumption there will always be cases in which the court would find the individual contributorily negligent) the defendant, knowing he would be able to escape liability in a certain fraction of the cases, would choose too low a level of care.

⁶² In *Carterville v. Cook*, 129 Ill. 152, 22 N.E. 14 (1889) the defendant maintained an

dant occurring after there has been time for them to absorb the significance of and react to a danger are more important to discourage than other acts, and the former but not the latter tend to result in exclusion.⁶³ Acts of parties other than the defendant that are reflexive, involuntary or compelled by another cannot, by definition, be controlled by imposing liability on the parties, and given such acts there is no special tendency toward exclusion of the defendant from the scope of liability.⁶⁴

F. *Proximity in Time and Space*

When there is a close temporal or physical connection between the injurer's action and the harm to the victim, the accident is more often found to be in the scope of liability than otherwise. Doubtless such proximity is associated on average with the factors that favor inclusion in the scope of liability under our theory. Specifically, when the injurer's action and the harm are closely rather than distantly linked in time and space, it is more likely that the effect of the injurer's action on accident losses is relatively high; that the type of accident was not difficult to imagine and therefore its probability was not underestimated (it was foreseeable); and that there were relatively few (or no) other important contributors to the harm. Consequently, that proximity in time and space serves as a prognosticator of inclusion in the scope of liability is not surprising.

But neither is it surprising that proximity has not been used with great success as a ruling criterion in decisions. What is only a correlate of the truly determinative factors of the scope of liability must have limited utility. Thus, when a defendant sends poisoned candy across the country or when he sets a spring gun that goes off many years later, he may be held responsible (his action has a sufficiently pronounced and foreseeable effect on expected accident losses) despite the lack of proximity.⁶⁵

G. *Sharing of Risk*

There is no reason to believe that the victim's ability to bear risk relative to the injurer's is related in a systematic way to any of the characteristics of

excavation close to a sidewalk. It was held that if another party negligently pushes the plaintiff into the excavation, the defendant is the proximate cause, except if the third party does so intentionally. See more generally the cases cited and discussion in Edgerton, *supra* note 10, at 364-67; see also Prosser, *supra* note 3, esp. at 274-75 & 282.

⁶³ See Prosser, *supra* note 3, at 288, where cases of the following type are cited. A defendant allows a child to come into possession of dynamite caps or some like object. If the child's parents do not discover the danger, even if because of their negligent inattention, the defendant is not excluded from the scope of liability when the parents are shown to be clearly aware of the danger.

⁶⁴ See the cases and discussion in Hart & Honoré, *supra* note 17, at 134; also Prosser, *supra* note 3, at 276.

⁶⁵ See citations and discussion in Edgerton, *supra* note 10, at 367.

accidents which were of interest to us in the model.⁶⁶ (Does the characteristic of causation in fact tell us anything about ability to bear risk?) Considerations of risk sharing therefore do not change the general conclusions drawn from the model.⁶⁷ However, when the relative ability of a party to bear risk or to cheaply insure is limited, it would be desirable for the scope of liability to be different from that suggested by our main argument. If, for example, injurers are better able to bear risk than victims in some class of accident situations, it would be advantageous to broaden somewhat the scope of liability.

H. *Extent of Liability*

Four reasons may be supplied for limiting the extent of liability. First, just as altering the scope of liability may help to spread risks, so may limiting the extent of liability. If accident losses are unusually high and a risk-averse and incompletely insured injurer is found negligent, it might be desirable to have him pay only partial damages.

Second, just as restricting the scope of liability reduces the likelihood of the injurer being discouraged from engaging in a socially worthwhile activity by a crushing burden of liability (see item 1(a) in the summary of Part V), so would limiting the extent of liability. Ideally, limiting the extent of liability to the injurer's incremental contribution to losses would eliminate the problem of the injurer inappropriately choosing not to engage in his activity. And, in practice, there is sometimes an attempt to limit the injurer's liability to his incremental effect on losses. It has been held that a prior condition (a disease, an injury) which limits the plaintiff's life expectancy will reduce the value of an action for wrongful death, that a defendant's responsibility for flood damage should be limited to the additional loss due to his negligence, and that a defendant's liability should extend only to the loss which would have occurred had the plaintiff not been intoxicated when he fell.⁶⁸ However, it is often difficult or impossible to determine such incremental losses. Losses may not be naturally divisible; and it often is difficult to unravel the effect of factors that have already come to bear (an individual who is exposed to a carcinogenic substance at his workplace develops lung cancer, but he also smokes five packs of Camels a day—what is the incremental contribu-

⁶⁶ The same might be said of the characteristic of unforeseeability. While it is true that, if an accident is unforeseeable, an injurer is less likely to have purchased adequate liability insurance coverage against it, the victim's purchase of first-party coverage is also less likely.

⁶⁷ As the reader is aware, these conclusions took into account only issues concerning incentives, for we assumed in the model that parties were risk neutral.

⁶⁸ See, respectively, *Fortner v. Koch*, 272 Mich. 273, 261 N.W. 762 (1935); *Baltimore & O.R. Co. v. Sulphur Spring Independent School District*, 96 Pa. 65 (1880); *O'Keefe v. Kansas City Western R. Co.*, 87 Kan. 322, 124 Pac. 416 (1912); generally, Hart & Honoré, *supra* note 17, 208-13; and Prosser, *supra* note 3, at 321.

tion of exposure to the carcinogen?); and it is sometimes hard to determine the effect of factors that would have come to bear later (X negligently drives his car over and kills a man sleeping in the middle of a country road—what would have happened to the man in X's absence?).

Third, just as restricting the scope of liability may be desirable on grounds of unforeseeability, so may be limiting the extent of liability.⁶⁹ If the probability of a high loss is underestimated, holding the injurer liable for the total loss rather than just for the foreseeable portion of it does not increase the incentive to take care. Of course, the same caveats mentioned in the discussion of unforeseeability apply here: It is often difficult to determine whether the magnitude of the total loss was unforeseeable; and limiting the extent of liability to the foreseeable harm may reduce the incentive to contemplate the full range of harm that might result from one's actions.

Last, it may be advantageous to limit the extent of liability because the potential for problems with multiple-injurers makes it difficult to place liability too far out in the (unending) sequence of events flowing from the defendant's act. Consider the following case.⁷⁰ The defendant negligently drove his taxi across the sidewalk and against the stoop of a house. Later, when employees of a towing truck company and several policemen were attempting to remove the taxi from its position on the stoop, a stone from high up on the house, which had been jarred loose by the original impact of the taxi, fell and injured the plaintiff who was standing nearby. Assume that there was a question as to the negligence of the towing company and policemen in the manner in which they attempted to remove the taxi, that the plaintiff perhaps should have hung back, that the stone that fell was already loose owing to the homeowner's not having kept his place in good repair. Rather than grapple with the problem of sorting out responsibility among the taxi driver, towing company, policemen, homeowner, and plaintiff, it might be easier for the court to limit the defendant's liability to the property damage done to the stoop.⁷¹

VII. OTHER VIEWS OF CAUSATION IN THE LAW OF TORTS

To relate the view of causation and the scope of liability presented here to other views, I shall divide them, admittedly with some arbitrariness, into four categories. The first category is comprised of systems of rules that, if adhered to, would have led the courts to the conclusions on causation

⁶⁹ What law is relevant is not entirely clear, but "the 'scope of the foreseeable risk' is becoming the criterion of what is 'proximate,'" Prosser, *supra* note 3, at 267.

⁷⁰ *In re Guardian Casualty Co.*, 253 App. Div. 360, 2 N.Y.S. 2d 232 (1938), and cited in Hart & Honoré, *supra* note 17, at 241.

⁷¹ In this example, unforeseeability is another reason (and, in fact, the reason given by the court) for limiting liability.

and liability actually reached. These systems of rules, among which those proposed by Beale and by Carpenter are probably best known, were never very successful.⁷² It was not difficult to adduce decisions that violated the rules. Moreover, the rules were phrased in terms of physical forces that are hard to interpret in many situations. The lack of success of the rules may be understood by analogy to our earlier comments on proximity in time and space. Just as proximity is only a correlate of the truly determinative factors of the scope of liability, these rules specify characteristics which are only correlates of those factors.

The second category, the "modern" one, contrasts greatly with the first. It is comprised of views which emphasize that while the scope of liability may be influenced by certain "objective" principles of causation, it also reflects the felt requirements of "policy"; and since the requirements of policy are so various, it will not be possible to find a system of rules that comprehends them. Views in the second category range from those of Andrews⁷³ and of Green,⁷⁴ who would have the courts rely almost entirely on intuition about policy, to those of Seavey and of Keeton,⁷⁵ who would have the courts employ the risk theory or, more generally, to those of Hart and Honoré who would have the courts use well-established causal notions as flexible, guiding principles in the determination of the scope of liability.

The third category of views is expressions of the so-called "adequacy theory" of von Kries and others of the German school.⁷⁶ According to this theory, the treatment of causation by the courts coincides with what would result from following two principles: first, for there to be liability, the injurer's lack of care should be a cause in fact of the accident in question;

⁷² Joseph H. Beale, *The Proximate Consequences of an Act*, 33 *Harv. L. Rev.* 633 (1920); and Charles E. Carpenter, *Workable Rules for Determining Proximate Cause*, 20 *Calif. L. Rev.* 229, 396, 471 (1932). See also the recent article, Richard A. Epstein, *A Theory of Strict Liability*, 2 *J. Legal Stud.* 151 (1973). This article proposes paradigmatic situations in which we would all agree that there should be liability; and, it is generally asserted that liability may be determined in a given instance by "closeness of fit" to one of the paradigmatic cases. For a criticism of this approach, see John Borgo, *Causal Paradigms in Tort Law*, 8 *J. Legal Stud.* 419 (1979).

⁷³ In his dissenting opinion in *Palsgraf v. Long Island R.R.*, *supra* note 49, Andrews said "What we do mean by the word 'proximate' is, that because of convenience, of public policy, of a rough sense of justice, the law arbitrarily declines to trace a series of events beyond a certain point. This is not logic. It is practical politics. . . . There are no fixed rules to govern our judgment." 248 *N.Y.* 353-54, 162 *N.E.* 103-04 (1928).

⁷⁴ Green, *supra* note 7, and the discussion of Green in Hart & Honoré, *supra* note 17, at chs. 4 & 10.

⁷⁵ Warren Seavey, *Mr. Justice Cardozo and the Law of Torts*, 39 *Colum. L. Rev.* 20, 29-39; 52 *Harv. L. Rev.* 372, 381-91; 48 *Yale L. J.* 390, 399-409 (1939); Keeton, *supra* note 52.

⁷⁶ I am relying here on the description given by Hart & Honoré, *supra* note 17, of the continental theories of causation. According to this description, *The Adequacy Theory*, developed by J. von Kries, *Die Prinzipien der Wahrscheinlichkeitsrechnung* (1886), and many successors, has enjoyed great success in German, Austrian, Swiss, and Italian civil courts.

and, second, his lack of care should have significantly, "adequately," raised the probability of the type of accident. The adequacy theory is of course consonant with much of what was said in this paper.

The remaining category of views derives from the instrumentalist approach, under which the choices that can be made over the scope of liability are related directly to well-specified social goals. Calabresi's view of causation⁷⁷ and the view of this paper are instrumentalist in nature. There is both an obvious similarity and an obvious difference between instrumentalist views and others. The similarity is that an appealing, relatively simple description of the determination of the scope of liability is sought under instrumentalist and other views (with the exception of those like Andrews' and Green's). And the difference is that only under the instrumentalist view is there a real effort to rationalize the description, to argue that the description is associated with choices about liability that best serve social goals. This difference is an advantage, for it allows us to think about causation and the scope of liability in a coherent and organized way; it allows us to see the sense behind and the connections between the many concepts and legal principles associated with causation.

However, there is a problem that may fairly be raised about the instrumentalist approach as it has been employed in the present context (and in most others). Questions about causation are to an important extent resolved by resort to intuitions about the justness of applying a rule of liability. In practice it is often asked when, according to common sense notions, liability of an injuring party for harm done ought to be contemplated; and it is not asked how liability would affect incentives or otherwise influence the attainment of certain basic social goals. Thus if the instrumentalist approach yields a successful explanatory theory, it must be that the intuitions about what is just comport with the application of a cost-benefit calculus in relation to the posited social goals.

Sometimes the validity of this claim is clear from the very discussion of what is just. Compare our explanation concerning causation in fact with Edgerton's:

⁷⁷ Calabresi, *supra* note 10. Calabresi stresses that many of the principles used to determine the scope and extent of liability provide desirable incentives to injurers to avoid accidents, at least when the form of deterrence implicit in the rules of liability is "market deterrence" rather than "specific deterrence." I have come to much the same conclusion in that the principles are most easily rationalized under the assumption that the form of liability is strict liability (Calabresi's market deterrence). But this article differs from Calabresi's in several respects. It formally deduced that something like the actual principles used are best, given a set of assumptions. It also placed more emphasis on the importance of administrative costs and identified for the first time, as a reason for restricting the scope of liability, the need to avoid imposing crushing liability. Additionally, it stressed somewhat different factors in discussing unforeseeability and several of the other subtopics in Part VII. However, as noted at the beginning, this article is definitely to be regarded as building on Calabresi's, and the reader is urged to turn to his article.

. . . our sense of justice demands the imposition of liability when the harm would not have happened but for the wrongful action of human beings, while it does not make the same demand when the harm would have been produced by a . . . natural force, if there had been no wrongful human action . . . if . . . D's wrongful act and a natural force were each sufficient to cause the harm, to hold D responsible and thereby discourage acts similar to his would tend less strongly to prevent the occurrence of similar harm, as it would have no effect on . . . natural forces.⁷⁸

Edgerton's argument about the justice of the requirement of causation is in its essence identical to the one given in this paper.

Other times the sense behind the claim is seen only after introspection about what lies behind our notion of justice and by careful comparison with the logic of the instrumentalist argument. For example, in *Berry* the court said: "That his [the streetcar driver's] speed brought him to the place of the accident at the moment of the accident was the merest chance . . . The same thing might as readily have happened to a car running slowly, or it might have been that a high speed alone would have carried him beyond the tree to a place of safety." Here, the court thought that it was not just to impose liability because speed did not alter the probability of the type of accident. I would guess that, to some degree, what was not articulated was a feeling that because speed does not alter the probability of such accidents, imposing liability for them would not alter streetcar drivers' behavior.⁷⁹ And this corresponds to the argument used here.

Yet, it also seems undeniable that the conceptions of when it is just to include a type of accident in the scope of liability have some life of their own, independent of any clear connection with postulated social goals through provision of incentives (or spreading of risks). Now there is no reason why, as a purely formal matter, the conceptions could not themselves be included among the social goals. However, their inclusion would detract from the analysis. More insight can be had by elaborating a parsimonious theory and by qualifying it at the end than by building into it all manner of assumption.

⁷⁸ Edgerton, *supra* note 10, at 347. There is much else in the second part of Edgerton's article that is strikingly similar to arguments given here.

⁷⁹ 191 Pa. 348, 43 A. 240 (1899). The skeptic may say that there is no liability *only* because slowing down would not have altered the probability of the type of accident, *not* because of an implication about the streetcar driver's incentives. In other words, we have ingrained in us the principle that it is not fair to hold a party liable for an accident when the injurer's action did not make the type of accident more likely. I might agree with the skeptic that this principle is ingrained in us and is sometimes employed without reference to instrumentalist goals. Thus I might agree that the instrumentalist analysis presented here is not, as positive theory, phenomenologically accurate. However, I am not overly disturbed about this. I would say it is incumbent on the skeptic to think about the origins of the principles of fairness under discussion here. Since they are often complex and highly refined, I do not regard the desire to explain them as unnecessarily reductionist; and I believe that if one attempted to explain them (say in the context of a model of childhood relationships with parents and peers), a theory not dissimilar from the one advanced here might emerge.

APPENDIX

As stated in Part III, we make the following assumptions. (a) There are two types of parties, injurers and victims, both of whom are risk neutral. (b) Each injurer must decide whether or not to engage in his activity and, if so, how much care to exercise; victims make no decisions. (c) Accident losses suffered by victims are assumed to be monetary or to have monetary equivalents, as is the benefit to an injurer from engaging in his activity and the disutility to him of taking care. (d) Social welfare is the sum of expected utilities; equivalently, it is the expected sum of monetary (or monetary equivalent) values, which means the sum of benefits from engaging in activities⁸⁰ less the costs of taking care less expected accident costs less administrative costs (to be defined later). Other assumptions to be used will be introduced as needed and, in any event, were described in Part III and in later parts of the text.

A. *The First-Best Solution to the Problem of Accidents*

Suppose first that the injurer is an individual who might engage in an activity. Let

v = benefit to the injurer of engaging in his activity.

This will be understood to be gross of the cost of taking care.

Now suppose the injurer to be a firm, that is, a seller of a good. In this case, it will be assumed for simplicity that, if the firm operates, it produces one unit of the good. Let

w = value of a unit of the good to consumers
 c = production cost (exclusive of care),

and then let

$$v = w - c. \quad (1)$$

Now define

- $p(s) > 0$ probability density of state of the world s
- $\ell(s) \geq 0$ loss to the victim if state s occurs and the injurer does *not* engage in his activity
- $x \geq 0$ level of care of the injurer (which is relevant only if he chooses to engage in his activity)
- $\ell(x, s) \geq 0$ loss to the victim if state s occurs, the injurer does engage in his activity, and x is the level of care.

The loss $\ell(x, s)$ as a function of x will be assumed to have one of two forms: (a) $\ell(x, s)$ is bounded (over x and s) and for each s it is a nonincreasing, differentiable, and convex nonnegative function of x . Thus, it might be identically zero (there is no loss), a positive constant (there is a positive loss that care doesn't affect), or positive at low levels of care but decreasing—possibly to zero—as care increases (care reduces the

⁸⁰ Victims' benefits are constant (since their actions are fixed by assumption) and may therefore be taken to be zero by appropriate choice of origin of their utility function.

magnitude of loss). (b) $\ell(x, s)$ is bounded (over x and s), and for each s it is zero for all but one value of x , at which it is positive. This would have been true in a continuous version of Example 2, concerning the streetcar, and of Example 5, concerning the car traveling down a country road (with regard to its striking down cattle); in both examples, an accident occurred only if speed (care) was such as to put the injurer in the precise position (the state of the world) of a suddenly appearing obstacle. It will be assumed that states associated with loss functions of type b can be grouped naturally into events e (that is, subsets of the set of all states) with the property that the expected loss conditional on the occurrence of such an event is a function of x of type a . For want of a better descriptive term, let us call such events "type- b events." In Example 5, the loss conditional on the type- b event "cattle cross the road at some point" is 100, independent of care. (If speed had increased the severity of an accident, then expected loss conditional on the event would have been a decreasing function of care.) Define $p(e)$ to be the probability density of e , $\ell(e)$ to be the expected loss conditional on e given that the injurer does not engage in his activity, and define $\ell(x, e)$ analogously.

If the injurer engages in his activity and chooses care level x , social welfare is

$$W(x) = v - x - \int_{[0,1]} \ell(x, t)p(t)dt, \quad (2)$$

where here and throughout the formal analysis t will be understood to stand either for a type- a state s or a type- b event e ; also, it will be assumed that t is a scalar in the unit interval. The functions ℓ and p will be assumed to be differentiable. Note that W is concave in x since ℓ is convex, and we will in fact assume that W is strictly concave. The first-best level of care given that the injurer engages in his activity will be denoted by x^* ; x^* is the nonnegative x that maximizes $W(x)$ or, equivalently, that minimizes $x + \int_{[0,1]} \ell(x, t)p(t)dt$.⁸¹ It will be assumed that x^* is positive; thus it is determined by the first-order condition

$$1 = - \int_{[0,1]} \ell_x(x, t)p(t)dt. \quad (3)$$

This says that the marginal cost of care must equal the expected marginal benefit (in terms of reduced accident losses) of care. From (3) it may be verified that x^* rises the more care reduces expected losses, the more probability mass is concentrated at states with more responsive loss functions, and so forth. Notice also that if $\ell(x, t)$ is a constant function of x over some range of t , then since $\ell_x(x, t) = 0$ for such t , the magnitude of losses over this range is irrelevant to the determination of x^* . Hence, in Example 5, as we knew had to be true, the mathematics indicate that the first-best speed is not influenced by the possibility of accidents with cattle but it is influenced by the possibility of accidents with tractors.

If the injurer does not engage in his activity, social welfare is simply

$$W = - \int_{[0,1]} \ell(t)p(t)dt. \quad (4)$$

⁸¹ Since ℓ is bounded over x and t and nonincreasing in x , it is easy to show that x^* exists.

Now it is socially desirable for the injurer to engage in his activity and to exercise care x^* when

$$W(x^*) \geq W \quad (5)$$

or when

$$v \geq x^* + \int_{[0,1]} (\ell(x^*, t) - \ell(t))p(t)dt. \quad (6)$$

This condition means that the gross benefit from the activity is greater than or equal to the cost of (first-best) care plus the expected *increment* in accident costs over their level were the injurer not to engage.

B. *The Achievable Solution to the Problem of Accidents Given Optimal Choice of the Scope of Liability under Strict Liability*

1. The solution in the absence of administrative costs. Let

$\mathcal{S} = \{t \mid \text{if } t \text{ occurs and there is an accident, the injurer is strictly liable for the victim's losses}\}.$

Given \mathcal{S} , the injurer's actions may be determined.

Assume first that the injurer is an individual who might engage in an activity. If he does not engage in the activity, his utility is 0. If he does engage in the activity and chooses level of care x , his expected utility is

$$Y(x, \mathcal{S}) = v - x - \int_{\mathcal{S}} \ell(x, t)p(t)dt. \quad (7)$$

Thus, if he does engage, he will choose x to maximize $Y(x, \mathcal{S})$. Since $-\ell$ is concave in x , so is Y ; and we will in fact assume that Y is strictly concave in x . Thus, the x that maximizes Y will be unique and will be denoted $x(\mathcal{S})$. To avoid having to consider uninteresting cases, we will assume that $x(\mathcal{S}) > 0$ for the \mathcal{S} of concern. Thus $x(\mathcal{S})$ is determined by the first-order condition

$$1 = - \int_{\mathcal{S}} \ell_x(x, t)p(t)dt, \quad (8)$$

which says that the marginal cost of care must equal the expected marginal reduction in liability payments. Note also that if $\mathcal{S}_1 \supset \mathcal{S}_2$, then $x(\mathcal{S}_1) \geq x(\mathcal{S}_2)$.⁸² The injurer will engage, choosing $x(\mathcal{S})$, if⁸³

$$Y(x(\mathcal{S}), \mathcal{S}) \geq 0. \quad (9)$$

Now assume that the injurer is a firm in a competitive industry in long-run (zero-

⁸² We know that $Y_x(x(\mathcal{S}_2), \mathcal{S}_2) = -1 - \int_{\mathcal{S}_2} \ell_x(x(\mathcal{S}_2), t)p(t)dt \leq -1 - \int_{\mathcal{S}_1} \ell_x(x(\mathcal{S}_2), t)p(t)dt = Y_x(x(\mathcal{S}_2), \mathcal{S}_1)$. And since $Y_x(x(\mathcal{S}_2), \mathcal{S}_2) = 0$, we have that $Y_x(x(\mathcal{S}_2), \mathcal{S}_1) \geq 0$. As Y is concave in x , we conclude that $x(\mathcal{S}_1) \geq x(\mathcal{S}_2)$.

⁸³ We will make the assumption that if equality holds in (9), the injurer will engage in the activity, and will make similar assumptions later without comment. Such assumptions are of course inessential.

profit) equilibrium and that victims are strangers. We wish to show that (9) is still the condition determining whether the injurer engages and that, if so, $x(\mathcal{S})$ is still the level of care. Define

$z =$ product price.

Since a firm takes price as given and attempts to maximize profits, it chooses x to maximize

$$z - c - x - \int_{\mathcal{S}} \ell(x, t) p(t) dt, \tag{10}$$

where, recall, c is production cost. Since this differs from $Y(x, \mathcal{S})$ by a constant, it follows that, if the firm operates, it will indeed choose $x(\mathcal{S})$. In addition, since profits are zero,

$$z = c + x(\mathcal{S}) + \int_{\mathcal{S}} \ell(x(\mathcal{S}), t) p(t) dt. \tag{11}$$

But, for consumers to buy the product, we must have

$$w \geq z, \tag{12}$$

where recall that w is the value of the good to consumers. Thus, the firm will operate and choose $x(\mathcal{S})$ when

$$w \geq c + x(\mathcal{S}) + \int_{\mathcal{S}} \ell(x(\mathcal{S}), t) p(t) dt \tag{13}$$

or, since $v = w - c$, when

$$Y(x(\mathcal{S}), \mathcal{S}) \geq 0, \tag{14}$$

which is (9).

If the injurer does engage in his activity, social welfare is (refer to (2)):

$$W(x(\mathcal{S})) = v - x(\mathcal{S}) - \int_{[0,1]} \ell(x(\mathcal{S}), t) p(t) dt, \tag{15}$$

and if he does not it is W (refer to (4)).

The problem to be solved is to choose \mathcal{S} to maximize social welfare, that is, to choose \mathcal{S} to maximize $\max(W(x(\mathcal{S})), W)$ subject to the constraint that (9) is the condition for the injurer to engage. We will assume that the problem has a solution \mathcal{S}^* and that \mathcal{S}^* may be taken to be a closed set.⁸⁴ The problem may be characterized as one of four types distinguished by the relationship between the first-best and achievable solutions. Only the last of the problem types will be of real analytical interest to us.

The problem is of the first type if under the first-best solution the injurer does not engage in his activity. In this case, under the achievable solution, the same will be true: Let $\mathcal{S} = [0, 1]$ so that $x(\mathcal{S}) = x^*$. (Compare (3) and (8)). Thus

⁸⁴ It will be seen that \mathcal{S}^* will not necessarily be unique (certainly it may be altered up to a set of measure zero).

$$Y(x(\mathcal{S}), \mathcal{S}) = v - x^* - \int_{[0,1]} \ell(x^*, t)p(t)dt = W(x^*). \quad (16)$$

But since under the first-best solution the injurer does not engage, $W(x^*) < W$, and since in addition $W \leq 0$, we conclude that $Y(x(\mathcal{S}), \mathcal{S}) < 0$, the injurer will not engage. Therefore, the first-best solution is indeed attained.

The problem is of the second type if under the first-best solution the injurer engages and under the achievable solution he does not. In such a problem, any \mathcal{S} under which the injurer would choose to engage in his activity is such that $W(x(\mathcal{S})) < W$. In other words, if the scope of liability is restricted enough to induce the injurer to engage in his activity, the level of care he would be led to take is so low that it would be socially undesirable for him to engage in it.

The problem is of the third type if under the first-best solution the injurer engages in his activity and the achievable solution coincides with the first best. One circumstance of this nature is when if $\mathcal{S} = [0, 1]$ (so that $x(\mathcal{S}) = x^*$), the injurer would choose to engage. Notice if $\mathcal{S} = [0, 1]$, for the injurer to choose to engage we must have

$$v \geq x^* + \int_{[0,1]} \ell(x^*, t)p(t)dt, \quad (17)$$

whereas (6) is the condition that under the first-best solution the injurer engages. Hence only if

$$v - x^* + \int_{[0,1]} \ell(x^*, t)p(t)dt \geq \int_{[0,1]} \ell(t)p(t)dt \quad (18)$$

would (6) imply (17). The interpretation of (18) is that surplus enjoyed by an injurer who engages and is strictly liable for all accident losses is at least as large as the expected losses that would be borne in his absence.

The problem is of the fourth type if under the first-best solution the injurer engages in his activity and under the achievable solution he also engages in it but does not take the first-best level of care. In other words, when the scope of liability is optimally restricted, the injurer has too little motive to take care; that is $x(\mathcal{S}^*) < x^*$.⁸⁵ This case, in which the constraint (9) is obviously binding, is the interesting case because it is the one in which the need to induce the injurer to engage in his activity necessitates a selection of the scope of liability which least reduces the incentive to take care.

PROPOSITION 1. Suppose that under the achievable solution the injurer engages in his activity and does not take the first-best level of care. Then there exists an \mathcal{S}^* obeying⁸⁶

⁸⁵ Since $\mathcal{S}^* \subset [0, 1]$, we know that $x(\mathcal{S}^*) \leq x^*$ by note 82 *supra*. And since, by hypothesis, $x(\mathcal{S}^*) \neq x^*$, it must be that $x(\mathcal{S}^*) < x^*$.

⁸⁶ It will be seen from the proof that \mathcal{S}^* is not unique. However, except for a set of measure zero, if an \mathcal{S}^* does not obey the form (19), then it is only because it also includes t such that $\ell(x(\mathcal{S}^*), t) = 0$; but these t must be interpreted as irrelevant, since if they occur there are no accident losses—there is no accident.

$$\mathcal{S}^* = \left\{ t \mid \frac{-\ell_x(x, t)}{\ell(x, t)} \geq k \right\}, \quad (19)$$

where $x = x(\mathcal{S}^*)$ and k is some positive constant.

Interpretation. (a) The determination of the scope of liability may be described as occurring in two stages. First, it must be found that the injurer's not having altered his level of care was a cause in fact of the accident. Second, it must be found that the accident situation was within a particular class; this class is comprised of all situations such that the marginal effect of care on accident losses, when normalized by the size of loss, is sufficiently great. To see that this interpretation follows from (19), note that since $k > 0$, it must be true that $-\ell_x(x, t) > 0$. In particular, suppose that what is usually meant by the injurer being a cause in fact is not true, that is, suppose that his level of care could not have affected accident losses in the circumstances. Then $\ell(x, t)$ is constant as x varies, so that $\ell_x(x, t) = 0$, which means that $t \notin \mathcal{S}^*$. The second claim is immediate from (19). (However, the level of care at which the partial derivative is evaluated is endogenous; thus specification of k alone does not determine \mathcal{S}^* .) The condition in (19) makes intuitive sense, for it is socially desirable to enhance the incentive to take care (thus high values of $-\ell_x$ work in favor of inclusion of t in \mathcal{S}^*) but socially undesirable to add to expected liability payments (thus high values of ℓ work against inclusion of t in \mathcal{S}^*).

(b) If the injurer's not having taken more care was a cause in fact of an accident, but given the circumstances (identified by a type- b event t), the relationship between care and the accident was "coincidental" (care did not affect expected losses given t), then the accident will not be in the scope of liability: What is meant by a coincidental relationship between care and accident losses is typified by *Berry*. Recall that given the type- b event $t =$ "a tree falls someplace on the route of the streetcar," the appropriate assumption—the assumption defining coincidental accidents—is that conditional expected losses $\ell(x, t)$ are independent of x . Thus $\ell_x(x, t) = 0$, so $t \notin \mathcal{S}^*$.

(c) The probability density of t is irrelevant to inclusion in the scope of liability: This is immediate from (19), which does not involve the density.

Proof. The argument is made in several steps.

(i) It may be assumed that $\ell(x(\mathcal{S}^*), t) > 0$ for $t \in \mathcal{S}^*$; thus the ratio of concern is well defined for the t we need to consider: If $\ell(x(\mathcal{S}^*), t) = 0$, then $\ell_x(x(\mathcal{S}^*), t) = 0$ (since ℓ is nonnegative). This means that if \mathcal{S}^* contains t with $\ell(x(\mathcal{S}^*), t) = 0$, we could remove such t without affecting the choice of x (as determined by (8)); moreover, removing such t would preserve (9). Hence, it may be assumed that \mathcal{S}^* contains no such t .

(ii) \mathcal{S}^* solves the problem

$$\text{Maximize } x(\mathcal{S}) \text{ over } \mathcal{S} \text{ satisfying (9)}. \quad (20)$$

To show this, note first that since it is optimal for the injurer to engage under the achievable solution, \mathcal{S}^* must in fact maximize $W(x(\mathcal{S}))$ over \mathcal{S} satisfying (9). Now suppose that some \mathcal{S}' and \mathcal{S}'' each satisfy (9). Then it suffices to demonstrate that $x(\mathcal{S}') > x(\mathcal{S}'')$ implies $W(x(\mathcal{S}')) > W(x(\mathcal{S}''))$. However, (see note 82) since $[0, 1] \supset \mathcal{S}'$, we have $x^* \geq x(\mathcal{S}')$, and similarly for \mathcal{S}'' . Hence, in fact, $x^* \geq x(\mathcal{S}') > x(\mathcal{S}'')$. Therefore, by strict concavity of W in x , the result follows.

(iii) \mathcal{S}^* solves the problem

$$\text{Maximize } - \int_{\mathcal{S}} \ell_x(x(\mathcal{S}^*), t) p(t) dt \text{ over } \mathcal{S} \text{ satisfying } Y(x(\mathcal{S}^*), \mathcal{S}) \geq 0. \quad (21)$$

To show this, suppose otherwise. Then, since by (8) $-\int_{\mathcal{S}^*} \ell(x(\mathcal{S}^*), t) p(t) dt = 1$, there must exist an $\mathcal{S}' \neq \mathcal{S}^*$ such that

$$-\int_{\mathcal{S}'} \ell_x(x(\mathcal{S}^*), t) p(t) dt > 1, \quad (22)$$

where

$$Y(x(\mathcal{S}^*), \mathcal{S}') \geq 0. \quad (23)$$

But (22) means that $x(\mathcal{S}') > x(\mathcal{S}^*)$ (by (8) and concavity of Y in x) and (23) means that $Y(x(\mathcal{S}'), \mathcal{S}') \geq 0$ (since $Y(x(\mathcal{S}'), \mathcal{S}') \geq Y(x, \mathcal{S}')$ for all x and, in particular, for $x = x(\mathcal{S}^*)$). However, this contradicts (ii).

(iv) There cannot exist disjoint nondegenerate intervals $[a, b]$ and $[c, d]$ such that $[a, b] \subset \mathcal{S}^*$, $[c, d] \cap \mathcal{S}^* = \emptyset$, and $-\ell_x(x(\mathcal{S}^*), t_1) / \ell(x(\mathcal{S}^*), t_1) < -\ell_x(x(\mathcal{S}^*), t_2) / \ell(x(\mathcal{S}^*), t_2)$ for all $t_1 \in [a, b]$, $t_2 \in [c, d]$: Assume otherwise. It is clear that the intervals can be chosen to be as small as required and such that

$$\int_{[a, b]} \ell(x(\mathcal{S}^*), t) p(t) dt = \int_{[c, d]} \ell(x(\mathcal{S}^*), t) p(t) dt. \quad (24)$$

Now modify \mathcal{S}^* by replacing $[a, b]$ with $[c, d]$ and call the new set \mathcal{S}^{**} . By (24), \mathcal{S}^{**} satisfies (9). But (24) also implies that for any $\alpha \in [a, b]$ and $\gamma \in [c, d]$

$$(b - a)\ell(x(\mathcal{S}^*), \alpha)p(\alpha) \approx (d - c)\ell(x(\mathcal{S}^*), \gamma)p(\gamma). \quad (25)$$

(Here \approx means approximately equal, and we rely on the fact that the intervals are chosen small and that ℓ and p are continuous.) Thus

$$(b - a) \approx \frac{(d - c)\ell(x(\mathcal{S}^*), \gamma)p(\gamma)}{\ell(x(\mathcal{S}^*), \alpha)p(\alpha)}. \quad (26)$$

Using this, we have

$$\begin{aligned} - \int_{[a, b]} \ell_x(x(\mathcal{S}^*), t) p(t) dt &\approx (a - b)\ell_x(x(\mathcal{S}^*), \alpha)p(\alpha) \\ &\approx \frac{(c - d)\ell(x(\mathcal{S}^*), \gamma)p(\gamma)\ell_x(x(\mathcal{S}^*), \alpha)p(\alpha)}{\ell(x(\mathcal{S}^*), \alpha)p(\alpha)} \\ &< \frac{(c - d)\ell(x(\mathcal{S}^*), \gamma)p(\gamma)\ell_x(x(\mathcal{S}^*), \gamma)}{\ell(x(\mathcal{S}^*), \gamma)} = \end{aligned} \quad (27)$$

$$(c - d)\ell_x(x(\mathcal{S}^*), \gamma)p(\gamma) \approx - \int_{[c, d]} \ell_x(x(\mathcal{S}^*), t) p(t) dt.$$

Therefore, $-\int_{\mathcal{S}^{**}} \ell_x(x(\mathcal{S}^*), t) p(t) dt > -\int_{\mathcal{S}^*} \ell_x(x(\mathcal{S}^*), t) p(t) dt$, which contradicts (iii).

(v) Completion of proof: It follows from (iv) (and from our assumption that \mathcal{S}^*

may be taken to be closed) that \mathcal{S}^* is of the required form, except that we have to show that $k > 0$. But if $k = 0$, then from (8) it is clear that $x(\mathcal{S}^*) = x([0,1]) = x^*$, contradicting the assumption that $x(\mathcal{S}^*) < x^*$. Q.E.D.

2. The solution when administrative costs are positive. Let

h = total administrative costs borne by all parties whenever an accident occurs
and $t \in \mathcal{S}$

$0 \leq \alpha \leq 1$ fraction of these costs borne by an injurer

Assume first that the injurer is an individual who might engage in an activity. If he does not engage in the activity, his utility is 0. If he does engage in it and chooses x , his expected utility is

$$Y(x, \mathcal{S}) = v - x - \int_{\mathcal{S}} \ell(x, t) p(t) dt - \alpha h G(x, \mathcal{S}), \quad (28)$$

where

$$G(x, \mathcal{S}) = Pr\{t | t \in \mathcal{S}, \ell(x, t) > 0\}. \quad (29)$$

We will assume that Y is strictly concave in x , we will denote by $x(\mathcal{S})$ the x that maximizes (28), and, to avoid uninteresting cases, we will suppose that $x(\mathcal{S}) > 0$. The first-order condition determining $x(\mathcal{S})$ is

$$1 = - \int_{\mathcal{S}} \ell_x(x, t) p(t) dt - \alpha h G_x(x, \mathcal{S}) \quad (30)$$

which differs from (8) only by the term $-\alpha h G_x(x, \mathcal{S})$; this term is nonnegative and reflects the expected decline in administrative costs borne by the injurer which would accompany an increase in care. The injurer will engage in his activity if (9) holds.

If the injurer is a firm, then, by a previous argument, (9) still determines the decision of the injurer whether to engage in his activity.

If the injurer does engage in his activity, social welfare is

$$W(x(\mathcal{S}), \mathcal{S}) = v - x(\mathcal{S}) - \int_{[0,1]} \ell(x(\mathcal{S}), t) p(t) dt - h G(x(\mathcal{S}), \mathcal{S}), \quad (31)$$

and if he does not engage in it, social welfare is W .

The problem is to choose \mathcal{S} to maximize social welfare; it is to choose \mathcal{S} to maximize $\max(W(x(\mathcal{S}), \mathcal{S}), W)$ subject to the constraint that (9) determines if the injurer engages. The reader, of course, realizes that the definitions of $x(\cdot)$ and Y are different from before. We will assume that a solution \mathcal{S}^* exists and that \mathcal{S}^* may be taken to be closed.

The problem may, as before, be characterized by type. The description of the first two types of problem is exactly as before. The third type is when under the first-best solution the injurer engages in his activity and under the achievable solution he also engages in it, but does not take the first-best level of care, and the constraint (9) is not binding. If (9) is not binding, there is no motive to reduce the scope of liability because of a concern that the injurer might not engage in his activity. Therefore, the only reason to reduce the scope of liability is to reduce administrative costs. The result for this type of problem is as follows.

PROPOSITION 2. Suppose that under the achievable solution the injurer does not take the first-best level of care and strictly prefers to engage in his activity. Then there exists an \mathcal{S}^* obeying

$$\mathcal{S}^* = \{t | -\ell_x(x, t) \geq k\}, \quad (32)$$

where $x = x(\mathcal{S}^*)$ and k is some positive constant.

Interpretation. The meaning of (32) is similar to that of (19).

Proof: The argument consists of several steps.

(i) It may be assumed that $\ell(x(\mathcal{S}^*), t) > 0$ for $t \in \mathcal{S}^*$: This is true by the logic of (i) in the proof of the previous proposition.

(ii) $\partial(v - x - \int_{[0,1]} \ell(x, t)p(t)dt)/\partial x > 0$ evaluated at $x = x(\mathcal{S}^*)$: By assumption, $x(\mathcal{S}^*) \neq x^*$ and it is easy to show $x(\mathcal{S}^*)$ cannot exceed x^* . Thus $x(\mathcal{S}^*) < x^*$ and the claim then follows since x^* maximizes W in (2) and W is strictly concave.

(iii) There cannot exist disjoint nondegenerate intervals $[a, b]$ and $[c, d]$ such that $[a, b] \subset \mathcal{S}^*$, $[c, d] \cap \mathcal{S}^* = \emptyset$, and $-\ell_x(x(\mathcal{S}^*), t_1) < -\ell_x(x(\mathcal{S}^*), t_2)$ for all $t_1 \in [a, b]$, $t_2 \in [c, d]$: Assume otherwise. It is then clear that the intervals may be chosen as small as required, to be of equal probability, and (by (i)) such that losses are bounded away from zero. Now modify \mathcal{S}^* by replacing $[a, b]$ with $[c, d]$ and call the new set \mathcal{S}^{**} . Since the intervals are small and (9) was assumed not to be binding, it is still satisfied by \mathcal{S}^{**} . Also, $x(\mathcal{S}^{**}) > x(\mathcal{S}^*)$. (Since ℓ is bounded away from zero in $[a, b]$ and $[c, d]$, $G_x(x, \mathcal{S}^*) \equiv G_x(x, \mathcal{S}^{**})$ in a neighborhood of $x(\mathcal{S}^*)$. From this and the contrapositive hypothesis, it is clear that at $x(\mathcal{S}^*)$ the right-hand-side of (30) exceeds 1 under \mathcal{S}^{**} , which, by strict concavity of Y in x , establishes the claim.) Thus, by (ii), the first three terms in social welfare rise. Also, since $x(\mathcal{S}^{**}) > x(\mathcal{S}^*)$ and (by construction of \mathcal{S}^{**}), $G(x(\mathcal{S}^*), \mathcal{S}^*) = G(x(\mathcal{S}^*), \mathcal{S}^{**})$, we must have $G(x(\mathcal{S}^{**}), \mathcal{S}^{**}) \leq G(x(\mathcal{S}^*), \mathcal{S}^*)$. Hence the last term in social welfare does not fall. Thus social welfare rises under \mathcal{S}^{**} , contradicting the optimality of \mathcal{S}^* .

(iv) Completion of proof: It follows from (iii) (and from the assumption that \mathcal{S}^* may be taken to be closed) that \mathcal{S}^* may be taken to be of the required form, except that we have to show that $k > 0$. But if $k = 0$, then from (30), $x(\mathcal{S}^*) \geq x^*$, and this contradicts (ii). Q.E.D.

The problem is of the fourth type if under the first-best solution the injurer engages in his activity and under the achievable solution he also does but the constraint (9) is binding. The interpretation is, therefore, that both the concern to induce the injurer to engage in his activity and the desire to save administrative costs are relevant to the determination of the scope of liability. I have not been able to simply characterize \mathcal{S}^* in this case.

C. *The Achievable Solution to the Problem of Accidents Given Optimal Choice of the Scope of Liability under the Negligence Rule*

1. The solution in the absence of administrative costs. Let

$\mathcal{S} = \{t | \text{if } t \text{ occurs and there is an accident, then if the injurer is found negligent, he must pay for the victim's losses}\}$.

\bar{x} = due care level.

It will be assumed that the courts observe and compare x to \bar{x} with complete accuracy. (Although this assumption was relaxed at the end of the analysis of Part VI, it will be maintained here.) Given \mathcal{S} , the injurer's actions may be determined.

Assume first that the injurer is an individual who might engage in an activity. If he does not engage in the activity, his utility is 0. If he engages in the activity and decides to act in a nonnegligent way, then clearly he will exercise the lowest possible nonnegligent level of care, namely \bar{x} . Thus, his utility will be

$$v - \bar{x}. \quad (33)$$

If he engages in the activity and decides to act in a negligent way—selects an $x < \bar{x}$ —then his expected utility will be

$$Y(x, \mathcal{S}) = v - x - \int_{\mathcal{S}} \ell(x, t)p(t)dt. \quad (34)$$

Thus, if he engages in the activity, for him to decide to act in a nonnegligent way, we must have

$$x + \int_{\mathcal{S}} \ell(x, t)p(t)dt \geq \bar{x} \quad \text{for all } x < \bar{x}; \quad (35)$$

the cost of care if he acts negligently plus expected liability payments must be at least as high as the cost of exercising due care. As in the informal analysis, we will consider only \bar{x} and \mathcal{S} such that if the injurer engages, he would decide to take due care (otherwise, the injurer would be, in effect, strictly liable, the situation already analyzed). That is, we will assume that \bar{x} and \mathcal{S} satisfy (35). Consequently, if the injurer engages, his utility will be $v - \bar{x}$. Therefore, he will engage in his activity if

$$v \geq \bar{x}. \quad (36)$$

Now assume that the injurer is a firm in a competitive industry in long-run equilibrium and that victims are strangers. Then the constraint (35) still guarantees that the firm would exercise due care were it to operate and (36) still determines whether it would decide to operate. This can be shown using an argument analogous to the one made in part B of the Appendix.

If the injurer does engage in his activity, social welfare is

$$W(\bar{x}) = v - \bar{x} - \int_{[0,1]} \ell(\bar{x}, t)p(t)dt, \quad (37)$$

and if he does not it is W .

The problem is to choose \mathcal{S} and \bar{x} to maximize social welfare, that is, to choose \mathcal{S} and \bar{x} to maximize $\max(W(\bar{x}), W)$ subject to the constraints (35) and (36) as the condition that determines whether the injurer engages. We will assume that the problem has a solution \bar{x}^* and \mathcal{S}^* , where \mathcal{S}^* may be taken to be closed.

The problem may be one of two types. It is of the first type if under the first-best solution the injurer does not engage in his activity. In this case, it is nevertheless possible (as should be clear from the discussion in the text) that under the achievable solution, the injurer engages.⁸⁷ If so, it will follow from the logic of the next para-

⁸⁷ See also Shavell, *supra* note 42.

graph that under the achievable solution, we may assume that $\mathcal{S} = [0, 1]$; there is no reason to restrict the scope of liability.

The problem is of the second type if under the first-best solution the injurer engages in his activity. In this case, given a reasonable assumption, under the achievable solution the same will be true: Let $\mathcal{S} = [0, 1]$ and $\bar{x} = x^*$: By definition of x^* ,

$$x + \int_{[0,1]} \ell(x, t)p(t)dt \cong x^* + \int_{[0,1]} \ell(x^*, t)p(t)dt \quad (38)$$

for all x . Thus, since

$$\int_{[0,1]} \ell(x^*, t)p(t)dt \cong 0,$$

(35) is satisfied, and the injurer would choose x^* if he decides to engage. Moreover, if we use (6) and make the assumption that for any x

$$\int_{[0,1]} \ell(x, t)p(t)dt \cong \int_{[0,1]} \ell(t)p(t)dt$$

(engaging in the activity can only increase expected accident losses), we have

$$v \cong x^* + \int_{[0,1]} [\ell(x^*, t) - \ell(t)]p(t)dt \cong x^* \quad (39)$$

so that from (36) the injurer would indeed decide to engage. Notice that the first-best solution is attained and that the scope of liability is not restricted.

2. Achievable solution when administrative costs are positive. Assume first that the injurer is an individual who might engage in an activity. If he does not engage in the activity, his utility is 0. If he does and acts in a nonnegligent way, his expected utility is

$$v - \bar{x} - \alpha hG(\bar{x}, \mathcal{S}) \quad (40)$$

and if he engages and acts in a negligent way, it is

$$v - x - \alpha hG(x, \mathcal{S}) - \int_{\mathcal{S}} \ell(x, t)p(t)dt. \quad (41)$$

We will continue to assume that \bar{x} and \mathcal{S} are such that were he to engage, he would decide to act in a nonnegligent way; thus

$$x + \int_{\mathcal{S}} \ell(x, t)p(t)dt + \alpha hG(x, \mathcal{S}) \cong \bar{x} + \alpha hG(\bar{x}, \mathcal{S}) \quad (42)$$

must hold for \mathcal{S} and $x < \bar{x}$. Consequently, he will engage in the activity if

$$v \cong \bar{x} + \alpha hG(\bar{x}, \mathcal{S}). \quad (43)$$

If the injurer is a firm (42) still determines whether the injurer would exercise due care if he engages and (43) determines whether he engages, by an argument parallel to that used earlier.

If the injurer does engage in his activity, social welfare is

$$W(\bar{x}, \mathcal{S}) = v - \bar{x} - \int_{[0,1]} \ell(\bar{x}, t)p(t)dt - hG(\bar{x}, \mathcal{S}) \tag{44}$$

and if he does not engage it is W . The problem is to choose \bar{x} and \mathcal{S} to maximize $\max(W(\bar{x}, \mathcal{S}), W)$ subject to (42) and to the condition that (43) determines whether the injurer engages in his activity.

The problem may be one of three types. It is of the first type if under the first-best solution the injurer does not engage in his activity. In this case, it is still possible that under the achievable solution the injurer engages in his activity. It is of the second type if under the first-best solution the injurer engages in his activity and under the achievable solution he does not. It is of the third type if under the first-best and achievable solutions, the injurer engages in his activity. This type of problem (and problems of the first type such that injurers engage) is the one of primary interest; the result about the scope of liability is as follows.

PROPOSITION 3. Suppose that under the achievable solution the injurer does not take the first-best level of care and chooses to engage in his activity. Then an \mathcal{S}^* obeys

$$\mathcal{S}^* = \{t | \ell(\bar{x}^*, t) \geq k\} \tag{45}$$

where k is some nonnegative constant.

Interpretation. The meaning of (45) is that the scope of liability is fully characterized by the size of accident loss. In particular, causation in fact, the ability of the injurer to affect accident losses, the probability density of t , and so forth are irrelevant to the determination of the scope of liability. The explanation for the result was given in Part VI.

In order to prove the result we will make the following assumption (though a somewhat weaker one would do).

$$\text{if } \ell(x, t) > \ell(x, t') \quad \text{for } t \neq t' \tag{46}$$

then

$$\ell(x', t) - \ell(x', t') \geq \ell(x, t) - \ell(x, t') \quad \text{for } x' < x.$$

This means that if at a given level of care x , the loss if t occurs exceeds the loss if t' occurs, then at any lower level of care x' , the loss if t occurs still exceeds the loss if t' occurs, and by at least as much.

Proof: (The proofs of (i) and (ii) and some other details of the argument will be omitted, as they are analogous to what was given in Propositions 1 and 2.)

(i) $\ell(\bar{x}^*, t) > 0$ for $t \in \mathcal{S}^*$

(ii) $\partial(v - \bar{x} - \int_{[0,1]} \ell(\bar{x}, t)p(t)dt) / \partial \bar{x} > 0$ evaluated at $\bar{x} = \bar{x}^*$:

(iii) There cannot exist disjoint nondegenerate intervals $[a, b]$ and $[c, d]$ such that $[a, b] \subset \mathcal{S}^*$, $[c, d] \cap \mathcal{S}^* = \emptyset$, and $\ell(\bar{x}^*, t_1) < \ell(\bar{x}^*, t_2)$ for all $t_1 \in [a, b]$, $t_2 \in [c, d]$: Assume otherwise. Then the intervals may be chosen as small as required, to be of equal probability, and such that losses are bounded away from zero. Now modify \mathcal{S}^* by replacing $[a, b]$ with $[c, d]$ and call the new set \mathcal{S}^{**} . Thus

$$\int_{\mathcal{S}^{**}} \ell(\bar{x}^*, t) p(t) dt \cong \int_{\mathcal{S}^*} \ell(\bar{x}^*, t) p(t) dt + \eta \quad (47)$$

for some positive η ; and by the assumption made preceding the proof, this must hold for $x < \bar{x}^*$ as well. Therefore,

$$x + \int_{\mathcal{S}^{**}} \ell(x, t) p(t) dt \cong x + \int_{\mathcal{S}^*} \ell(x, t) p(t) dt + \eta \quad (48)$$

for all $x \leq \bar{x}^*$. Hence, (see (42)) under \mathcal{S}^{**} the due care level can be raised somewhat to an $\bar{x}^{**} > \bar{x}^*$. By (ii), this raises the first three terms in social welfare. Also,

$$\begin{aligned} G(\bar{x}^{**}, \mathcal{S}^{**}) &= Pr\{t | t \in \mathcal{S}^{**}, \ell(\bar{x}^{**}, t) > 0\} \leq Pr\{t | t \in \mathcal{S}^{**}, \ell(\bar{x}^*, t) > 0\} \\ &= Pr\{t | t \in \mathcal{S}^*, \ell(\bar{x}^*, t) > 0\} \\ &= G(\bar{x}^*, \mathcal{S}^*), \end{aligned} \quad (49)$$

so the last term in social welfare cannot have fallen. This means social welfare is higher given \bar{x}^{**} and \mathcal{S}^{**} , a contradiction.

(iv) Completion of proof: It follows from (i) and (iii) (and the assumption that we may take \mathcal{S}^* to be closed) that (45) holds. Q.E.D.