ISSN 1936-5349 (print)
ISSN 1936-5357 (online)

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

AN ALTERNATIVE TO THE BASIC CAUSAL REQUIREMENT FOR LIABILITY UNDER THE NEGLIGENCE RULE

Steven Shavell

Discussion Paper No. 1100

03/2023

Harvard Law School
Cambridge, MA 02138

An Alternative to the Basic Causal Requirement for Liability under the Negligence Rule

Steven Shavell*

The primary causal requirement that must be met for a negligent party to be held liable for a harm is a demonstration that the harm would not have occurred if the party had not been negligent. Thus, for a speeding driver to be found liable for harm done in a car accident, it must be shown that the accident would not have happened if the driver had driven at a reasonable speed. The main point made here is that this basic causal requirement may be difficult to satisfy and hence may interfere with the discouragement of negligence. Therefore, an alternative and usually easier-to-meet causal requirement is proposed—that the harm would not have occurred if the party had not been engaged in his activity (if the driver had not been driving).

1. Introduction

In this article I will suggest that the principal causal requirement for a finding of liability under the negligence rule may needlessly hamper deterrence of negligent conduct; and in light of that, I will consider a possibly advantageous alternative requirement.

At present, a negligent party cannot be held liable for a harm unless the party's *negligence* was a cause in fact of the harm—meaning that the harm would not have occurred if the party's conduct had not been negligent, that is, if the party had instead exercised reasonable care.[1] Suppose that a driver negligently speeds on a city street and strikes a pedestrian on a crosswalk, whereas if the driver had been traveling at the speed limit, he would have been able to stop short of the

* Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School, and Research Associate, National Bureau of Economic Research. I thank John Goldberg, Marcel Kahan, Louis Kaplow, A. Mitchell Polinsky, and David Rosenberg for advice and comments on this article, Daniel Belgrad, Tuhin Chakraborty, Heather Pincus, and Lisa Wang for able research assistance, and the John M. Olin Center for Law, Economics, and Business at Harvard University for research support.

[1] See Restatement, Third, Torts: Liability for Physical and Emotional Harm (hereafter "Restatement Third") § 26 Factual Cause, stating that "Tortious conduct must be a factual cause of harm for liability to be imposed. Conduct is a factual cause of harm when the harm would not have occurred absent the conduct." See also Becht and Miller (1961) and Dobbs, Hayden, and Bublick (2016) ch.14. For the most part, I will use the terms cause in fact, factual cause, or sometimes simply cause in this article instead of other commonly employed synonyms, notably, actual cause, but for cause, necessary cause, and cause sine qua non. Additionally, I note that Restatement Third, § 27 Multiple Sufficient Causes, accords wider meaning to the term factual cause. To avoid distracting issues, however, I will restrict attention to the definition of factual cause of § 26. Finally, I observe that in other Common Law countries and in civil law countries generally, negligence must also be shown to be a cause in fact of harm for a party to be held liable for it; see, for example, Infantino and Zervogianni (2017). Thus the § 26 cause-in-fact requirement for liability under the negligence rule appears to be a universal feature of tort law.

crosswalk and prevent the accident. Here the driver's negligence would be a cause in fact of the accident and the driver could be held liable for it.

It can readily happen, however, that a speeding driver brings about an accident even though his excessive speed would not be a cause in fact of that event. Suppose that a speeding driver encounters an area of sheer ice and slides on it into a parked car; and suppose also that if the driver had not been speeding, he would still have lost control of his vehicle on the ice and collided with the parked car. In this instance, the driver's negligent speeding would not be a cause in fact of the damage to the parked car and he could not be held liable for it.

Cases resembling the preceding illustration, in which a negligent party might be able to avoid liability because the party's negligence might not be found to be a cause in fact of a harm, are routinely addressed by courts. In *Joshi v. Providence Health System*, 149 P.3d 1164 (Or. 2006), for example, the claimed negligence of physicians and a hospital in initially failing to diagnose a patient's stroke was held not to be a cause in fact of his death because his stroke was so serious that he was likely to have died from it even if it had been correctly diagnosed; in *O'Grady v. State of Hawaii*, 398 P.3d 625 (Hawaii 2017), the asserted breach by Hawaii of its duty to provide the public a reasonably safe highway system was not held to be a cause in fact of motorists' injuries from a falling boulder because it was unclear whether the existence of a satisfactory highway safety program would have averted the particular rockfall in question; and in *Fedorczyk v. Carribean Cruise Lines Ltd.*, 82 F.3d 69, 73 (3d Cir. 1996), the alleged negligence of the defendant cruise lines in not having affixed an adequate number of adhesive safety strips to the bottom surface of a stateroom bathtub was not held to be a cause in fact of a slip and fall injury to a passenger because it was not demonstrated that a properly higher number of adhesive strips would have prevented her specific accident.[2]

There thus appears to be a real chance that negligent parties will be able to circumvent liability for harm because of the requirement that their negligence must be found its cause in fact. This

[2] Let me add several other typical illustrations. In *Morris v. National Seating & Mobility*, WL 2343020 (Texas 2019), the alleged negligence of the defendant in maintaining a wheelchair was not found to be a cause in fact of its collapse and injury to its occupant because the occupant's weight might have led to its failure in any event; in *June v. Union Carbide Corp.*, 577 F.3d 1234 (10th Cir. 2009), the defendant had engaged in uranium milling operations and allegedly exposed individuals living in a town to radioactive materials, but the plaintiffs were unable to establish that they would not have suffered from their illnesses for other reasons in the absence of their exposure; and in *Yearty v. Scott Holder Enterprises*, 349 Ga. App. 718 (Georgia 2019), the defendant was claimed to have negligently installed a smoke alarm that failed to warn the plaintiff of a kitchen fire, resulting in a burn injury, whereas evidence suggested that she might still have been burned if the smoke alarm had functioned properly. See generally Restatement Third § 26 Comment for well-organized and valuable critical observations about factual causation and illustrations of it, and see Reporters' Note for extensive discussion of literature on, and cases involving, factual causation.

possibility adds to existing general reasons that negligent parties might escape liability[3] and implies that there may be a social need to bolster the threat of liability that they face.

A straightforward way to enhance the likelihood of liability for negligence is to enlarge the set of situations under which a negligent party would be found to have been a cause in fact of harm. That would be accomplished by a proposal that I will advance here: permit causation in fact to be demonstrated by a showing that a negligent party's *activity* was a cause in fact of harm—that if the negligent party had not been engaged in his activity when the accident occurred, the accident would not have happened.[4] In the example of the speeding driver who slid on ice into a parked car, we know that the driver's *speed* would not be a cause in fact of his accident. Yet the driver's activity—namely*, driving his car*—would be a cause in fact of his accident because, if he had not been driving on the road, he could not have slid into the parked car.[5]

I will argue in Section 2 below that because use of the activity-based cause-in-fact criterion will tend to raise the probability of liability for negligence, it can promote two socially desirable outcomes. First, it can improve deterrence of negligent conduct. And second, it can alleviate a problem of socially excessive engagement in dangerous activities. (This problem may arise especially because a party who is negligent thereby renders an activity that would not normally be particularly dangerous into a relatively dangerous one.) For purposes of clarity and precision, I develop these points in the world of a stylized model of liability for accidents under the negligence rule.[6]

In Section 3 I state and elaborate on my proposal, discussing among other matters (i) how the proposed requirement would be applied in practice,[7] (ii) why the proposed requirement would be satisfied more often than the present requirement, (iii) why litigation costs would tend to fall under the proposed requirement, (iv) the mistaken notion that the proposed requirement could lead to excessive liability because it is activity-based, and (v) a skeptical reaction to the view that it is unfair and hence undesirable to impose liability on a negligent party unless his negligence was a cause in fact of harm.

---

[3] Notably, the evidence needed to prove negligence might be lacking, the identity of an injurer might not be known, or an injurer's assets might be too low to make suit worthwhile.

[4] As will be discussed in Section 3, the proposal would also permit causation to be proved as it is now, through a showing that negligence was a cause in fact of harm.

[5] Nevertheless, in some circumstances another car could have slid on the ice into the parked car. See Section 3.2 for a discussion of this and like possibilities.

[6] I do so using numerical examples that should be accessible to all readers. Algebraic verification of the claims I make is provided in an appendix to the article.

[7] This will include attention to causal requirements going beyond those of causation in fact.

Before proceeding, let me mention writing relevant to this article. As readers will be aware, there exists a rich and expansive literature on causation from philosophical[8] and legal perspectives.[9] What is most germane to my effort here is writing on causal requirements for liability from the point of view of their functionality—notably, how causal requirements may foster deterrence of dangerous behavior. In this regard, Calabresi (1975) is of importance because, to my knowledge, his is the earliest article that examines causation in tort law in self-consciously economic terms. The central argument of his article applies mainly to strict liability. Specifically, it is that imposing damages on a party when his activity was a cause in fact of harm implies that the party will pay for the harms that the activity imposes on society; and thus the party will be motivated to reduce dangers in a socially desirable manner.[10] However, Calabresi sees no comparable role for the use of the cause-in-fact requirement under the negligence rule.[11] In Shavell (1980), I first formalized notions of causation in the standard model of accidents and liability and showed that if liability is restricted to situations in which a party's *activity* was a cause in fact of harm, individuals would be led to exercise socially desirable levels of care under both the strict liability rule and the negligence rule.[12] Additionally, in a significant article, Kahan (1989) demonstrated that if liability is further restricted to cases in which a party's *negligence* was a cause in fact of harm, individuals would be led to take desirable levels of care under the negligence rule.[13]

2. Cause-in-Fact Requirements in a Model of Liability for Negligence

2.1 Assumptions and framework of analysis

I will examine here a stylized model in order to obtain a clear understanding of the argument described in the Introduction. Consideration of such a model will allow us to abstract from a

[8] Philosophical contributions of primary significance are Hume (1739-1740) book I part III, Hume (1748) § IV part II, § V part I, § 7 parts I-II, and Mill (1843) book III. These are helpfully reviewed in Hart and Honoré (1985) ch. 1 and in Wright and Puppe (2016), which also contains a useful discussion of contemporary philosophical considerations of causation.

[9] The legal literature addressing causation and tort law deals not only with cause in fact but also with proximate cause, the umbrella term often employed to refer to further causal requirements for liability. On legal literature and reference to cases bearing on proximate causation and the scope of liability, see Restatement Third ch. 6 and Dobbs, Hayden, and Bublick (2016) ch. 15.

[10] See his discussion of what he discusses as market or general deterrence, especially on pp. 84-86.

[11] See his discussion on pp. 79-81.

[12] This conclusion rests on the assumption that law enforcement would be perfect—that suit would be brought against a party if he would be found liable. See also Shavell (1987) ch. 5 and Shavell (2004) pp. 249-253. As will be seen here in Section 2.4, the assumption of perfect law enforcement will be relaxed.

[13] Other economically oriented articles (the last being critical of the approach) on causation in tort law include Ben-Shahar (2009), Cooter (1987), Fennell (2022), Gilead and Green (2017), Grady (1984), Hylton (2014), Landes and Posner (1983), Porat (2011), and Wright (1985).

welter of complications that would divert our attention from the main incentives associated with the use of a cause-in-fact requirement for finding liability under the negligence rule. The analysis of the model will be informal in the sense that numerical examples will be employed to illustrate claims; the generality of the claims—that they are true in the world of the model not only in the numerical examples presented in the text—will as I noted be formally demonstrated in an Appendix.

We will focus on a single person, a potential injurer, who will choose among several actions; a state of the world will then occur; and a consequence will result, determined by the action and the state of the world.[14] The following table represents the model that we will consider.

Table 1. Consequences Given States of the World and Actions

Actions of person	State S (Safe road conditions)	State R (Risky road conditions—slippery leaves present)	State D (Dangerous road conditions—sheer ice present)
Person does not engage in activity (does not drive on road)	No harm	No harm	No harm
Person engages in activity and takes care (drives & does not speed)	No harm	No harm (driver able to maintain control of his car and avoid sliding into parked car)	Harm (driver slides into parked car)
Person engages in activity and does not take care (drives & speeds)	No harm	Harm (driver slides into parked car)	Harm (driver slides into parked car)

As can be seen, the model involves three possible actions—not engaging in an activity, engaging in it and taking care, or engaging in it and not taking care; and two consequences—no harm or harm. Shown in parentheses is an interpretation of the model to which I will usually refer for concreteness. Under this interpretation, mentioned in the Introduction, we see that when road conditions are safe, the person's action does not affect the consequence—harm does not occur; when road conditions are risky on account of slippery leaves, the person's action can affect the consequence—if he drives and does not speed, he will be able to slow down and avoid sliding into a parked car, whereas if he speeds, he will unavoidably slide into the parked car; and when

[14] This framework of actions, states of the world, and resulting consequences is that of decision theory. See, for example, Raiffa (1968) and Savage (1972) 6-17, 20-21, 27-30, 56-68, two classic sources on the subject.

the road is dangerous due to a coating of sheer ice, the person's speed will not matter to the consequence—he will slide into the parked car whether or not he speeds.[15] The model thus may be viewed as an exemplar of situations in which engaging in an activity generates a risk of harm and when failure to exercise care might be the cause in fact of harm (in State R) and yet might not be the cause in fact of harm (in State D).

Let me now continue with the description of the model. Assume that the probabilities of the three states in Table 1 are these.

Table 2. States of the World and Probabilities

State of the world	Probability
S (Safe)	30%
R (Risky)	50%
D (Dangerous)	20%

Let us also suppose the following.

Table 3. Variables and Their Magnitudes

Variable	Magnitude
Value of activity (driving)	60
Cost of care (refraining from speeding)	25
Possible harm (damage to parked car)	100

At this point, the main elements of the model have been described. The probabilities of the three states, the value of the activity, the cost of care, and the potential harm define a particular version of the model.[16]

---

[15] It may be helpful to describe another interpretation of the model of Table 1. Suppose a party's activity is the transport of oil by supertanker and the exercise of care corresponds to the use of an experienced crew. State S is safe in that the supertanker is not near dangerous submerged rocks; State R is risky in that, although the weather is normal, the supertanker comes into the vicinity of the dangerous rocks; and State D is dangerous in that the supertanker comes into the area of the rocks during a fierce storm. Harm is puncturing the hull of the supertanker on the rocks resulting in an oil spill. Finally, when the supertanker is not near the dangerous rocks, harm would not occur; when the supertanker is in the vicinity of the rocks and the crew is experienced, no harm will occur, whereas if the crew is not experienced, harm will occur; and if the supertanker is in the area of the rocks during a fierce storm, harm will occur whether or not the crew is experienced.

[16] To be clear, the structure of Table 1 is a fixed part of the model.

We will now proceed to consider two issues: first, the socially desirable behavior of the party who might generate harm; and second, the self-interested behavior of the party given each of two cause-in-fact requirements for liability under the negligence rule.

2.2 Socially desirable behavior

Let us adopt a simple utilitarian definition of social welfare, namely, the value of the activity if undertaken less the cost of care if exercised and less also any harm that comes about. I will sometimes refer to the cost of care and to harm as social costs of an activity.

To determine socially desirable behavior—that which maximizes social welfare—it is convenient to consider initially the behavior that will be best under the presumption that the party undertakes his activity. Then we can say whether it will be desirable for the party to engage in the activity by comparing its value to him to the social costs that the activity would generate.

Thus, let us begin in our example by assuming that the party engages in the activity of driving and ask whether it would be desirable for him to take care not to speed. If he speeds, the parked car will be damaged in states R and D, whose probabilities of 50% and 20% add to 70%, so that 70%×100 or 70 will be the expected harm.[17] If he does not speed, entailing a cost of care of 25, harm will occur only in state D, for harm will be avoided in state R. Hence expected harm will be only 20%×100 or 20, and thus social costs will be 25 + 20 or 45. Thus social costs will be lower when the driver does not speed than when he does.

A different way of expressing the conclusion that refraining from speeding is socially worthwhile is that its cost of 25 is less than the reduction in expected harm of 50 that it brings about.

Now that we know that ideal behavior is for the party not to speed if he drives and that driving and not speeding would generate social costs of 45, we can see that social welfare will be enhanced if he drives—for its value to him of 60 exceeds 45.

From the foregoing discussion, it should be clear that the general conclusion describing socially ideal behavior in the model (that is, allowing the probabilities of the states and the values of the variables to be different) is this: (a) the activity should be undertaken if and only if its value to the party would exceed the resulting minimized social costs (cost of care if that is desirable plus expected harm); and if the activity is undertaken, (b) care should be exercised provided that its

[17] "Expected" harm refers to probability-discounted harm, the amount that would occur on average were the party repeatedly to find himself in the situation described.

cost is less than the reduction in expected harm that it would accomplish.[18] In our version of the model, care was desirable to exercise, but if that were not the case—say the cost of care were 80—then the social costs of the activity would be 70%×100 or 70, implying that it would not be desirable to engage in the activity because we assumed the value of the activity was 60.

2.3 Behavior under two different cause-in-fact requirements for liability under the negligence rule—when law enforcement is perfect

We define negligence in the model to be the failure to take care when the exercise of care would have been socially desirable; equivalently, negligence is failure to take care when the cost of care is less than the expected reduction in losses that it would engender. This corresponds to the usual conception of negligence in the law.[19] In our version of the model, speeding would be negligent, for the cost of refraining from speeding, 25, is less than the reduction of 50 in expected harm that this form of care brings about on account of State R.

We will assume that for a negligent party to be held liable for a harm, his negligence and the occurrence of harm is not enough. His conduct must also have been found to be a cause in fact of the harm. In this regard, we will examine each of the two definitions of cause in fact mentioned in the Introduction.

We will also assume for now that law enforcement is perfect—meaning that if a negligent party brought about harm under an applicable definition of cause in fact, then he would definitely pay damages of 100 for the damage to the parked car.

Behavior under the negligence-based cause-in-fact requirement. As discussed in the Introduction, the present definition under tort law of conduct having been a cause in fact of a harm under the negligence rule is that a party's negligence brought about the harm. Hence, we will say here that a party's *negligence was a cause in fact of a harm* if the harm would have been prevented had the party exercised appropriate care.

In our model, a driver's speeding would be a cause in fact of damage to the parked car if that harm have been avoided had the driver not speeded. This would be true in State R involving slippery leaves, for if the driver had speeded in that state, he would have slid into the parked car, whereas if he had not speeded, he would have been able to control his car so as to avoid sliding into the parked car. However, speeding would not be a cause in fact of damage to the parked car in State D involving sheer ice, for the driver would have slid into the parked car on account of that condition whether or not he had been speeding.

[18] See Proposition 1 in the Appendix.

[19] See, for example, Restatement Third § 3 cmt. e and Reporters' Note cmt. d.

It follows that if the driver decides to speed, he will be liable for the damage of 100 to the parked car only in State R, so that his expected liability would be 50%×100 or 50. However, if he decides not to speed and thus would avoid liability for negligence, his only expense would be the cost of 25 of refraining from speeding. Hence, the driver will be induced to take adequate care under the negligence rule when he would be liable for negligence only when his negligence would be a cause in fact of harm.[20]

The foregoing conclusion should not be surprising, for under the negligence-related cause-in-fact requirement, the party's exposure to liability is in State R—which is precisely when the exercise of care would prevent hitting the parked car and thus permit the party to save liability expenses. In State D, by contrast, the exercise of care would not save liability expenses.

Let us next examine the other aspect of behavior in which we are interested, namely, whether the party would choose to engage in his activity in the first place. The answer in our version of the model is that the party would do so: Because the party would be led to refrain from speeding by the threat of liability in State R, he would not bear liability. Accordingly, his only expense would be the cost of care, 25. And because his benefit from the activity of driving is 60, he would undertake it. The outcome is thus socially desirable—the party chooses to engage in the activity of driving and exercises care when doing so.

A qualification to this conclusion about the social desirability of behavior in our version of the model should be mentioned, however. Namely, it is possible in other versions of the model that the party would choose to engage in his activity even though that would be socially undesirable. Suppose, for instance, that the value of the activity to the party were 35. Then he would engage in it, for 35 exceeds the cost of care of 25. But since the total social cost of engaging in the activity of driving is the cost of care of 25 *plus* the expected harm in State D, which is 20%×100 or 20, the social costs add to 45 as had been observed. Hence, it would not be socially desirable for the party to engage in the activity of driving. This problem is an example of a general defect in the negligence rule—that it can lead to excessive participation in potentially harmful activities because it does not impose on parties liability that reflects harms caused by activities despite the exercise of proper care.[21]

---

[20] See Proposition 2 in the Appendix. This result, that the negligence-based cause in fact requirement induces non-negligent behavior (when law enforcement is perfect), is, as I noted in the Introduction, originally demonstrated in Kahan (1989).

[21] See Proposition 4(c). The general point about the defect in the negligence rule is first developed in Shavell (1980); see also Shavell (1987) ch. 2, and Restatement Third § 20 cmt. b, Reporters' Note cmts. b and k.

To summarize our conclusions about behavior under the negligence-based cause-in-fact requirement: The party will be induced to engage in the activity when that would be socially desirable, and when doing so he would be led to take proper care; but the party might also be induced to engage in the activity when that would not be socially desirable, although in this case too he would be led to take proper care.

Behavior under the activity-based cause-in-fact requirement. Consider next the other definition of cause in fact of harm. In particular, we will say that a party's *activity* was *a cause in fact of a harm* if the harm would not have occurred had the party not participated in his activity.

In our model, it is apparent that a party's driving would be a cause in fact of damage to the parked car in both State R and State D—for had the party not been driving, his car would not have been on the road so as to slide into the parked car. Consequently, the activity-based cause-in-fact requirement would result in liability for a speeding driver in both States R and D, meaning that his expected liability if he was speeding would be 70%×100 = 70. Since the cost of care is 25, we see again that the driver would be led to take care and refrain from speeding.[22]

We also see that a party would choose to engage in the activity of driving, as his value would be 60 and his cost of care would be 25. Moreover, the qualification that was made above (vis a vis the negligence-based cause-in-fact requirement) that his choice whether to engage in driving might be socially excessive still holds.

Accordingly, the summary statement that I made about the negligence-based cause-in-fact requirement would obtain here. Nevertheless, there is a difference between the two cause-in-fact requirements: *the threat of liability for negligence would be greater under the activity-based requirement*; it would be expected damages of 70 as opposed to 50. But the smaller expected damages threat of 50 is still sufficient to induce care-taking under the assumption that law enforcement is perfect. As we will now verify, however, our conclusions will be different when law enforcement is not perfect.

2.4 Behavior under different cause-in-fact requirements for liability under the negligence rule—when law enforcement is imperfect

In reality, the expected liability faced by parties who behave negligently will often be less than what the law would require because of imperfect law enforcement. As I mentioned in the Introduction, negligent parties might not be found negligent even though they were truly negligent. Several prominent reasons for this phenomenon are that the identity of the injurer might not be known to the victim, that evidence needed to prove negligence might be lacking, and that the injurer's assets might be too low to make full payment of a judgment possible or to

[22] See Proposition 2 in the Appendix. That the activity-based cause-in fact-requirement induces non-negligent behavior (when law enforcement is perfect) is first shown in Shavell (1980b); see also Shavell (1987) ch. 5.

make the bringing of suit worthwhile. To capture the phenomenon of imperfect law enforcement in this section, I will assume for simplicity that there is only a probability that a party will be found liable when he ought to be found liable.

Behavior under the negligence-based cause-in-fact requirement. As discussed in Section 2.3, a negligent party will be liable under this cause-in-fact requirement only in State R, when not speeding would have prevented damage to the parked car. Thus, the party's expected liability if he was speeding would have been 50%×100 or 50.

Now, however, our assumption is that a party who ought to bear liability will bear it only with a probability. Let us say that this probability of bearing liability is q. For instance, if the probability q is 80%, a driver who speeds would face expected liability of 80%×50 = 40, rather than 50. However, in this instance, the driver would still have an incentive not to speed, since 40 exceeds the cost of care 25. But if the probability q is low enough, the driver will not be induced to refrain from speeding. The probability at which the party would the party would be at the cusp of deciding not to speed is that for which q×50 = 25, which is to say 50%. In other words, if q is less than 50%, the party would decide to be negligent and speed, and if *q* exceeds 50% the party would have sufficient incentive not to speed.[23]

Let us next consider the issue of participation in the activity. In this regard, we need to consider that in our particular version of the model, *if the party is negligent, then his activity will become socially undesirable*. The reason is that if the party is negligent, he will cause harm in States R and D, so that the expected harm he generates will be 70. But the value of the activity to him is only 60. Thus, social welfare will fall by 10 if he engages in the activity of driving.

When I asked two paragraphs above whether the party would be led to act negligently for various probabilities q of being found liable, I did not remark on whether a negligent party would choose to engage in his activity of driving. But it is clear that a negligent party would do so. The reason is that because he has *chosen* to be negligent, it must be that his expected liability was less than his cost of care, 25.[24] However, the benefit to a party from driving is 60. Because the benefit of 60 exceeds 25 and thus his expected liability, we know that he will engage in the activity. Accordingly, we conclude that for probabilities q of liability less than 50%, not only will a party decide to be negligent, he will also socially undesirably engage in the activity of driving.

Behavior under the activity-based cause-in-fact requirement. Our discussion of this cause-in-fact requirement parallels that of the negligence-based cause-in-fact requirement. Specifically, a negligent party whose driving was a cause-in-fact of harm will be liable in States R and D, so that his expected liability would be 70 under perfect law enforcement. Hence, his expected

[23] See Proposition 5 in the Appendix.

[24] For example, if q is 40%, the driver's expected liability if he is negligent would be 40%×50 = 20. As 20 is less than the cost of care of 25, the driver would indeed elect to be negligent.

liability for negligence under imperfect enforcement will be q×70. Accordingly, the critical q below which the party would not be motivated to take care and not speed is such that q×70 = 25, implying that q is 35.7%. Hence, for q below 35.7% the party would speed and for q above this probability, the party would be refrain from speeding.[25]

With regard to participation in the activity, our conclusion is that when q is below 35.7%, the party will be negligent but will still engage in his activity. That will be socially undesirable for the reasons given above in regard to the negligence-based cause-in-fact requirement.

Comparison of the two cause-in-fact requirements. As we have seen in our version of the model, *negligence is deterred more often under the negligence rule when it is accompanied by the activity-based cause-in-fact requirement than when it is accompanied by the negligence-based cause-in-fact requirement*. Specifically, for probabilities between 35.7% and 50%, negligence is deterred under the activity-based cause-in-fact requirement but not under the negligence-based cause-in-fact requirement.

Moreover, because in our version of the model participation in the activity is socially undesirable when parties act negligently, it follows that *inappropriate participation in the activity occurs more often under the negligence-based cause-in-fact requirement than under the activity-based cause-in-fact requirement* (again for the probabilities between 35.7% and 50%).[26]

3. A Proposal in Light of the Foregoing Analysis

The advantages of the activity-based cause-in-fact requirement that were identified in the model of the preceding section suggest that the requirement might be beneficial for the legal system to employ in practice. I now state a proposal for doing so and consider a number of issues bearing on it.

3.1 A proposal for use of the activity-based cause-in-fact requirement

As I indicated in the Introduction, the proposal I am advancing is that *the cause-in-fact requirement for liability under the negligence rule can be satisfied either by showing that the*

[25] See Proposition 6 in the Appendix.

[26] The point just made rests on the assumption that the value of the activity was 60, namely, less than the expected harm due to the activity if the driver speeds, 70. If the value of the activity exceeded 70, such as 75, then the activity would be socially desirable despite the occurrence of negligence. Notwithstanding such possibilities, it is true that the only kind of difference that use of the negligence-based cause-in-fact requirement rather than the activity-based cause-in-fact requirement can make to outcomes is a socially undesirable one; see Propositions 7 and 8 in the Appendix.

defendant's participation in his activity was a cause in fact of harm or else, as now, by demonstrating that the defendant's negligence was a cause in fact of harm.[27]

It should be noted that the recommendation that the cause-in-fact requirement can be satisfied by establishing that a harm would not have occurred if the defendant had not engaged in his activity is equivalent to a recommendation that *the defendant be held strictly liable for harm provided that he was found negligent*.[28] I will remark on this equivalence below because it can be a useful lens for viewing the proposal.[29]

3.2 Application of the activity-based cause-in-fact requirement

The question whether a party's activity was a cause in fact of a harm for which he was prima facie responsible I believe would often be easily answered in the affirmative. In our example of the driver whose car slid on ice into a parked car, the activity-based causal question was whether the parked car would have been damaged if the defendant had not been driving his car. And the answer to that question was self-evident in the world of the model, for there the presumption was that if the defendant driver was not on the road, the parked car could not have been damaged. Would the answer be different in reality? It is conceivable that if the defendant had not been driving his car, another driver might have slid on the ice into the parked car,[30] meaning that the activity of the defendant would not be a cause of the harm to it. However, as a practical matter, my surmise is that such a possibility would be unlikely to rise to a real level of concern in the kind of case at issue, especially for evidentiary reasons.[31] Thus, my expectation is that the defendant's activity would be viewed by a court as a clear cause of the damage to the parked car.

I come to similar conclusions about most cases that I have read. Consider two mentioned in the Introduction. In *Fedorczyk v. Carribean Cruise Lines Ltd.*, the activity of Caribbean Cruise Lines was providing cruises to customers. Hence, the activity-based causal question would be whether Fedorczyk would have been injured in a bathtub accident if Carribean had not been engaged in

[27] The proposal would leave unchanged causal requirements for liability concerning multiple sufficient causes and proximate cause; see Restatement Third § 27 and ch. 6.

[28] See Restatement Third §§ 3, 26. In particular, §26 cmt. h makes it clear that the cause-in-fact question under strict liability is whether the harm would have occurred absent the defendant's *activity*.

[29] See Sections 3.2 and 3.4.

[30] This possibility can be recognized as causal in terms of the general formal definition of causation that makes use of the framework of decision theory. In particular, suppose the description of the state of the world is expanded to include not only the condition of the road (icy when the accident occurred) and the presence of the parked car, but also the location and paths of travel of other cars. Then whether the parked car would have been struck by another car if the defendant had not been driving would be answered by the question (given the state of the world in its expanded form) would the parked car have been struck by another car.

[31] It would be difficult for the court to obtain much information about the location and the paths of other vehicles on the road as well as other potentially relevant matters, such as when the parked car might have been moved and changes in weather conditions that might have altered the slipperyness of the ice. Without reliable evidence on such issues, a court seeking to predict the probability of another vehicle hitting the parked car would quickly find itself in a realm of unacceptable conjecture.

its business and thus if Fedorczyk had not been a passenger on Carribean's vessel. Under that hypothesis, it is improbable that she would have been involved in a different bathtub accident, say at her home or on a different vessel;[32] and thus it would almost automatically be concluded that Caribbean's activity was a cause in fact of her accident. Likewise in *O'Grady v. State of Hawaii*, the activity of the defendant was supplying appropriately safe roads to drivers. Thus, the activity-based causation question would be whether, in the absence of the availability of the road where a rockfall occurred, the O'Grady plaintiffs would have been injured by a rockfall on a different road. The answer to this question would almost surely be no.[33]

As illustrated by the preceding cases, my sense is that the determination of activity-based causation would be immediate or almost so in a substantial majority of cases and result in a finding that the defendant's activity was a cause of the harm. However, I am not claiming that counterexamples would be rare. For instance, in *June v. Union Carbide*, the defendant had allegedly negligently exposed plaintiffs to radioactive materials, but the thyroid diseases that some of them developed could have been due to factors other than their dangerous exposure to Union Carbide's materials.[34] I therefore believe that a court might well not have found Union Carbide's activity to have been a cause in fact of plaintiffs' harms.

In the cases that I have mentioned in this section I have made predictions about how courts would apply the activity-based cause-in-fact concept (the courts were, of course, actually applying the negligence-based causal criterion). We do, however, have direct knowledge of how the courts have applied activity-based causation from strict liability cases. And in these cases, what we generally see is essentially what I have suggested above that we would encounter. For instance, in the well-known strict liability case of *Siegler v. Kuhlman*, Wash., 502 P.2d 1181 (Wash. 1972), the defendant was the driver of a gasoline tanker truck that spilled gasoline onto a highway, leading to an explosion and fire that burned to death a driver who found herself at the wrong place at the wrong time. What the court said about causation in fact was brief and to the point, that "searing flames . . . engulfed her car" and that "[t]he result of the explosion is clear…." Presumably the reason that the court felt no need to explicitly address activity-based causation is that it was manifest that if the defendant's truck had not been on the road, the fire and the plaintiff's death would not have occurred.[35] Similarly, in other strict liability cases, involving blasting, explosion

[32] If, as I suppose, the chances of a bathtub accident were numerically low per use by Fedorczyk, so too would be the probability that she would have experienced this kind of accident had she used a bathtub elsewhere for some relevant small number of times. Yet it would have to shown that she would have faced a probability of a bathtub accident elsewhere of over 50% for it to be held that her actual accident was not the cause of her injury.

[33] Only under highly unusual circumstances would this question even have relevance. Suppose, say, that evidence demonstrated that the O'Grady plaintiffs would have driven on a nearby road if the road in question did not exist or had been closed to traffic. Under this assumption, how would a rockfall come about? One would have to imagine that rocks that could fall were present on this other road and that some set of circumstances (perhaps seismic activity) made a rockfall likely there at the time the actual accident occurred. Even then, the odds that the plaintiffs would have been at the right point on the road to be struck by the rocks would be miniscule.

[34] See note 2.

[35] *Siegler v. Kuhlman*, 1182.

of stored ammunition, fireworks displays, and the like, the defendant's activity is commonly treated as a cause of harm in a largely proforma manner because specific consideration is unnecessary.[36]

Still, strict liability cases in which issues of causation in fact require serious consideration do occur. For instance, in *Dyer v. Maine Drilling & Blasting, Inc.*, 984 A.2d 210 (Me. 2009), the defendant engaged in significant blasting operations over almost a year near the plaintiff's home. She claimed that vibrations from the blasting caused extensive damage to her dwelling, including a several inch subsidence of the structure and sagging of the first floor. However, other possible causes of the damage to her property were mentioned by an expert, including earth pressure, ground water, and changes in temperature. Thus, the matter of causation in fact was unclear.[37]

Finally, let me remark on the point made in Section 3.1 that liability for activity-based causation is equivalent to imposition of strict liability on parties who have been found negligent. This observation raises the question whether activity-based causation could lead to imposition of liability for *any* harm due to a negligent party's activity. For instance, suppose that we agree that Carribean Cruise Lines was negligent in failing to affix a proper number of adhesive strips to the bathtub bottom in Fedorczyk's stateroom; and suppose too that, contrary to the facts in the case, Fedorczyk did not slip in her bathtub but did fall and injure herself when walking along the hallway on the way to her stateroom. Under activity-based causation, Fedorczyk could apparently make a claim for damages: she could assert that if Carribean had not engaged in the activity of operating its cruise lines, she could not have fallen in the hallway of its vessel; and since Carribean was negligent in having provided her an unsafe bathtub, she could collect damages for her injury. Is it true that Fedorczyck could succeed in her claim about a fall in the hallway under the proposal made in this article?

The answer to this question is that I envision the proposal to be accompanied by causal requirements going beyond those of factual causation—to those described as coming under the

---

[36] For example, Restatement Third § 20 cmt. e remarks that "blasting causes harm essentially on its own, without meaningful contribution from the conduct of the victim or of any other actors." See for illustration *Balding v. D. B. Stutsman, Inc.*, 54 Cal. Rptr. 717 (Ct. App. 1966) in which blasting activity resulted in an injury due to a stone projectile; *Yukon Equipment v. Fireman's Fund Ins. Co.,* 585 P.2d 1206 (Alaska 1978) concerning the detonation of a storage magazine for explosives; *Klein v. Pyrodyne Corp.*, 810 P.2d 917 (Wash. 1991) relating to a fireworks display; *Zero Wholesale Gas Col, Inc., v. Stroud*, 571 S.W. 2d 74 (Ark. 1978) concerning a gas explosion; and *National Steel Service Center v. Gibbons*, 319 N.W. 2d 269 (Iowa 1982) involving a gas explosion.

[37] *Dyer v. Maine Drilling & Blasting, Inc.*, 214. Another strict liability case involving uncertainty over causation is *Caporale v. C. W. Blakeslee and Sons, Inc.*, 175 A. 2d 561 (Conn. 1961). This case concerned harm to a building alleged to be due to vibrations from pile driving operations, but another possible cause was vibrations from heavy truck traffic on a neaby road.

heading of scope of liability or proximate cause.[38] In particular, I assume that one of the most important such requirements applies—that liability be limited to "those harms that made the actor's conduct tortious," sometimes referred to as "harms within the scope of the risk."[39] In the Fedorczyck hypothetical under discussion, this causal requirement would immediately bar damages for Fedorczyck's injury in the hallway. That is because the harms that make failure to affix adhesive strips to a bathtub bottom negligent concern slips and falls in bathtubs and are unrelated to falls in hallways. Thus, the within-the-scope-of-the-risk causal criterion prevents parties from making a wide range of claims for harms whose likelihood is unconnected to the negligent conduct at issue in a case.

3.3 Why the proposal would lead to more frequent findings of causation-in-fact

The central argument for the claim that the proposal will elevate the frequency of findings of causation in fact is simply that the proposal would provide plaintiffs with a second avenue for establishing that causal requirement. This additional opportunity can only add to the number of findings of causation and would tend to do so measurably.

Specifically, findings of causation will increase relative to their level now when (1) activity-based causation can be shown in a negligence case, whereas (2) negligence-based causation cannot be established, either because it is not true or because proof of it is too expensive or impossible to adduce. I would expect this double circumstance to hold with real frequency.

On one hand, as was discussed in Section 3.2, activity-based causation will often be essentially a manifest fact and will be readily demonstrable. Hence, I had suggested as plausible that activity-based causation could be fairly easily established in more than half of all negligence cases.

On the other hand, parties often fail to show negligence-based causation, as I observed in the Introduction. This was so in the cases I mentioned there, and treatises on torts and the Restatement Third describe a multiplicity of examples.[40] It seems reasonable to estimate that in about half of these cases in which negligence-based causation cannot be shown, activity-based causation could be shown, given my assumption that such causation can be demonstrated in at least half of all cases. Accordingly, the number of additional cases in which the causation-in-fact requirement would increase under the proposal would be substantial.

[38] See Restatement Third ch. 6 and Dobbs, Hayden, and Bublick (2016) ch. 15; these sources prefer the term scope of liability to proximate cause. I will not address in this article the functional desirability of scope of liability requirements. I instead am saying that I assume that the factual causation proposal would be accompanied by the various scope of liability requirements that the courts generally employ.

[39] See Restatement Third §29; and see also Dobbs, Hayden, and Bublick (2016) §§15.1-15..

[40] See Restatement Third §26 and Dobbs, Hayden, and Bublick (2016) §§14.1-14.5.

3.4 The proposal would often lead to a reduction in litigation costs

The primary argument for this claim is again that the proposal gives plaintiffs an additional means of demonstrating causation in fact. Because plaintiffs will be motivated to choose the less expensive method of proving causation, we can infer that when they choose the activity-based route, they will usually be lowering their litigation costs.

Furthermore, plaintiffs' savings could be significant. That is because the costs of showing activity-based causation are likely to be low, as I argued in Section 3.2.

Moreover, demonstrating negligence-based causation can easily become complicated. This can be appreciated by considering a few examples. In the illustration that I offered in the Introduction of a speeding driver who hits a pedestrian at a crosswalk, the determination of causation could be difficult even though the accident is mundane in character. The question at issue would be whether the driver would have been able to avoid striking the pedestrian had the driver been traveling at the speed limit. This inquiry would focus on evidence that could be obtained about two critical variables: the distance between the driver and the pedestrian when the driver was first was able to see him on the crosswalk; and the minimum distance the driver would need to bring his car to a halt if he were obeying the speed limit. If the latter distance was less than the former, the accident could in theory have been avoided had the driver not exceeded the speed limit.[41] Some reflection on the details of such an inquiry reveals, however, that ascertaining the two distances might not be at all straightforward.[42] In *Fedorczck v. Carribean Cruise Lines Ltd.*, the negligence-based causation question was whether, if the number of adhesive strips on the

[41] Suppose that the driver first saw the pedestrian when the driver was 50 feet away and that he could have brought his vehicle to a stop were he moving at the 25 mph speed limit within 35 feet. Then in theory the accident could have been avoided since if he had been able to stop in 35 feet, there would still have been 15 feet separating his car from the pedestrian on the crosswalk.

[42] How would information on the the distance between the driver and the pedestrian when the driver was first able to see him on the crosswalk be determined? Let us assume that this would require a statement from a disinterested witness. Would such a witness be able to identify, much less remember, what the position of the defendant driver's car was when the driver could first see the pedestrian on the crosswalk? How would the witness know what that point was? Among other matters, this would require the witness to have observed when the pedestrian first stepped onto the crosswalk (if the pedestrian did that when the car was close to the crosswalk the situation would be quite different from that if the pedestrian had stepped onto the crosswalk much earlier). Moreover, the witness would have to take into account factors affecting the visibility of the pedestrian to the driver, including the time of day and the clothing of the pedestrian. Now let us consider briefly the other distance determination, that of the stopping distance of the driver's vehicle if traveling at the 25 mph speed limit. This distance would depend on assessments of the reasonable ability of a driver to react to a suddenly perceived danger, the braking characteristics of his car, the nature of his tires, and whether he might have gone into a skid. In other words, many variables would be involved in the assessment. Altogether, it is easy to imagine real debate and uncertainty about the evidence on the two distances needed to determine whether the accident would have been avoided had the driver not been speeding.

bottom of the bathtub had been greater, equal to some number deemed proper, Fedorczck would not have fallen and suffered an injury. In examining this matter, the court addressed issues including whether Fedorczck's feet would likely have been primarily resting on adhesive strips if the number of them was seven rather than four (the actual number) and the degree to which her use of soap or bath oils might have affected the performance of the strips. The evidence bearing on these questions was limited, rendering judgments about them difficult to make.[43] Likewise, the nature of the inquiry about negligence-based causation in *O'Grady v. State of Hawaii* was challenging. There the question was, if Hawaii had developed a reasonable safety program to avert accidents due to rockfalls, what its nature would have been; and given that conjectural exercise, how would the danger of the rockfall that did occur have been assessed and would it have been prevented?[44]

In the previous paragraph, I am not meaning to assert that the costs of determining negligence-based causation generally tend to be high—I would expect them usually to be low or modest because causation will be obvious.[45] My surmise is instead that the costs *can* be high in a not insubstantial percentage of cases. This would be significant because it is what would create the opportunity for the proposal to reduce litigation costs.

3.5 The mistaken notion that use of activity-based causation in fact could lead to excessive liability

By excessive liability, I refer to liability burdens that could exert an undesirable chilling effect on participation in socially beneficial activities. I suspect that some readers of this article will be concerned that the proposal might lead to such undue liability because, as I have observed, the proposal can be regarded as a species of strict liability, albeit imposed only on negligent parties.

---

[43] Opinion at 71, 74-75.

[44] A circuit court had regarded a reasonable rockfall safety program as one that would be primarily concerned with remediation of dangerous sites (opinion at 630-631); ostensibly the program would undertake projects to move rocks and earth. Hence, to prove that a remediation program would have prevented the fall of the boulder that harmed the plaintiffs, the circuit court found that plaintiffs should have submitted evidence about the cost of preventing the boulder from falling and whether that cost would have been considered worth bearing by those administering a remediation program. The Supreme Court of Hawaii observed, however, that a reasonable safety program might have contemplated not only engaging in remediation projects but also monitoring rockfall dangers and the possibility of posting warnings to drivers or of closing roads. According to this broader view of a satisfactory safety program, its existence might have prevented the rockfall accident that occurred (opinion at 637-638) and the Supreme Court thus remanded the case to the circuit court for further consideration.

[45] Suppose in the example of the driver and the pedestrian that it was apparent that the driver was able to see a slow moving pedestrian on the crosswalk on a bright and sunny day from a distance of over 100 feet and the minimum stopping distance at the speed limit would have been 25 feet. See also, for example, Malone (1958) at 68-72 and at 71 his view that "In the mine run of cases the facts are so clear that the issue of causal relation need not even be submitted to the jury."

I do not believe, however, that there is a meaningful danger of excessive liability associated with the proposal. First, as I noted at the end of Section 3.2, the scope of liability facing negligent parties would be limited to harms within the risks motivating negligence. Second, parties can in principle endeavor to protect themselves from liability by exercising adequate care.[46] Third, an overarching reason that readers ought not fear excessive liability is that as a general matter, the use of strict liability[47] should not lead to socially undesirable suppression of activities. Indeed, and on the contrary, strict liability is *needed* to meliorate the problem of socially excessive engagement in potentially dangerous activities that occur under the negligence rule. Although this contention is antithetical to what many in the legal community believe to be true of strict liability,[48] the claim is a well-known conclusion of economic analysis of tort law based on the logic of internalization that many regard as intellectually compelling.[49]

3.6 The view that it would be unfair to impose liability for negligent conduct that was not a cause in fact of harm

If we put to the side the claimed merits of the proposal with regard to standard functional goals of tort law—in this article deterrence of dangerous conduct and reduction of litigation cost—it is natural to ask whether the proposal would offend the felt demands of individuals for the law to display fairness toward those who come before it. Here that general question devolves into whether individuals' sense of fairness would include the to-me rather refined principle that liability for harm in a negligence case should be imposed *only* when a person's negligence was a cause in fact of a harm rather than *also* when a person's activity was its cause in fact.

---

[46] The protection will, of course, be limited by the possibility of legal error.

[47] By strict liability I mean to refer to strict liability accompanied by possible defenses of contributory or comparative negligence.

[48] Skepticism of the legal community about the propriety of use of strict liability can be inferred from the raw fact that the domain of use of strict liability is greatly limited. See Restatement Third, ch. 4. Moreover, discussions about strict liability generally presume that the negligence rule ought to be the norm and thus adopt the perspective that the possible warrant of strict liability is narrow. That is made plain, for example, in the series of comments in Restatement Third § 20.

[49] See Shavell (1980a). The kernel of the logic underlying the claim is that there is a fundamental defect in the negligence rule: parties who behave in a non-negligent manner by definition escape liability for harms that their activities still generate. Hence, they will tend to engage too often in activities to the degree that risk exists despite the exercise of due care. (Nevertheless, the conclusions drawn in the article are qualified in various ways within the world of the stylized model studied.) A separate advantage of strict liability should also be emphasized—that under strict liability parties will be motivated to choose *all* dimensions of care so as to appropriately lower risks, but under the negligence rule parties will not be led to exercise care in the many dimensions of behavior that courts cannot effectively police (for instance, how often drivers check their rear view and side mirrors). For additional development of these points, see Shavell (1987) chs. 2-3 and Shavell (2018). Finally, it is worth recognizing that in the past, before liability insurance became widely available in the early decades of the 1900s, concerns about an undesirable chilling effect of strict liability had some basis.

In this regard, let us briefly reconsider the illustration I offered of the speeding driver who slides on ice into a parked car and would have done so as well had he not been speeding—implying that he would be absolved of liability under our law today because his speeding was not a cause of the damage to the car. On learning about a judicial decision against liability, I can imagine a statement of consternation made by the plaintiff to the defendant along the following lines. "Well, my car is a complete loss, which is more than a little bothersome because I have no collision insurance coverage on it. I would have thought I'd win my suit because you were obviously entirely responsible for the accident. My car was parked, you chose to drive when the roads had icy spots, and you caused the accident that totaled my car. To top that off, you were speeding at the time. For you to go scot-free because, as my lawyer tells me, your speeding was somehow not a legally recognized cause of losses seems to me like a technical excuse. Why should I care that your car might have slid into mine even if you were obeying the speed limit? For me, the court's decision is almost inexplicable and adds insult to injury."

The words of the plaintiff in the foregoing conversation I believe capture views that would be held by many individuals in society—apart from those in the legal community. That community is largely wedded to the specific requirement of negligence-based factual causation, because they view it as natural and just.[50] On what underlying moral rationale would defenders of this traditional factual causation requirement accord it precedence over the activity-based requirement, and how would these defenders regard the consequentialist problems with the traditional requirement adduced here?

[50] Virtually all of the writing on factual causation with which I am familiar endorses the use of the negligence-based cause-in-fact requirement and finds that it is in basic accord with the notion that moral responsibility for a harm must lie in the defendant's having caused the harm in some intuitive sense. See generally Restatement Third § 26 and Dobbs, Hayden, and Bublick (2016) ch. 14. For other writing on the cause-in-fact requirement, see, for example, Becht and Miller (1961), Edgerton (1924), Hart and Honoré (1985), Malone (1958), and Smith (1911). The main exception to the general approval of the requirement concerns multiple sufficient causes, as I had mentioned earlier (note 1), on which there is agreement that the negligence-based causation requirement should be relaxed, as the law does.

References

Becht, Arno C., and Frank W. Miller. 1961. *The Test of Factual Causation in Negligence and Strict Liability Cases*. Washington University: St. Louis.

Ben-Shahar, Omri. 2009. Causation and Foreseeability. *Tort Law and Economics* (Michael Faure, ed.), ch. 3, 83-108.

Calabresi, Guido. 1975. Concerning Cause and the Law of Torts: An Essay for Harry Kalven, Jr. 43 *University of Chicago Law Review* 69-108.

Cooter, Robert. 1987. Torts as the Union of Liberty and Efficiency: An Essay on Causation. 63 *Chicago-Kent Law Review* 523-551.

Dobbs, Dan B., Paul T. Hayden, and Ellen M. Bublick. 2016. *Hornbook on Torts*, 2nd ed. West Academic Publishing: St. Paul.

Edgerton, Henry W. 1924. Legal Cause. 72 *University of Pennsylvania Law Review* 211-244, 343-375.

Fennell, Lee Anne. 2022. The Gun and the Paperweight: Risk Control Services and Disservices. 15 *Journal of Tort Law* 67–107.

Gilead, Israel and Michael D. Green. 2017. Positive Externalities and the Economics of Proximate Cause. 74 *Washington and Lee Law Review* 1517-1566.

Grady, Mark F. 1984. Proximate Cause and the Law of Negligence. 69 *Iowa Law Review* 363-449.

Hart, H. L. A., and Tony Honoré. 1985. *Causation in the Law*, 2nd ed. Clarendon Press: Oxford.

Hume, David. 1739-1740. *A Treatise of Human Nature*. Andrew Millar: London.

Hume, David. 1748. *An Enquiry Concerning Human Understanding*. Andrew Millar: London.

Hylton, Keith N. 2014. Information and Causation in Tort Law: Generalizing the Learned Hand Test for Causation Cases. 7 *Journal of Tort Law* 35-64.

Infantino, Marta, and Eleni Zervogianni. 2017. *Causation in European Tort Law*. Cambridge University Press: Cambridge.

Kahan, Marcel. 1989. Causation and Incentives to Take Care under the Negligence Rule. 18 *Journal of Legal Studies* 427-447.

Landes, William, and Richard Posner. 1983. Causation in Tort Law: An Economic Approach. 12 *Journal of Legal Studies* 109-134.

Malone, Wex S. 1958. Ruminations on Cause in Fact. 9 *Stanford Law Review* 60-99.

Mill, John Stuart. 1843. *A System of Logic Ratiocinative and Inductive.* London.

Porat, Ariel. 2011. Misalignments in Tort Law. 121 *Yale Law Journal* 82-141.

Raiffa, Howard. 1968. *Decision Analysis*. Addison-Wesley: Reading, Ma.

Savage, Leonard J. 1972. *The Foundations of Statistics* (second revised ed.). Dover: New York.

Shavell, Steven. 1980a. Strict Liability versus Negligence. 9 *Journal of Legal Studies* 1-25.

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

Shavell, Steven. 1987. *Economic Analysis of Accident Law*. Harvard University Press: Cambridge.

Shavell, Steven. 2004. *Foundations of Economic Analysis of Law*. Harvard University Press: Cambridge.

Shavell, Steven. 2018. The Mistaken Restriction of Strict Liability to Uncommon Activities. 10 *Journal of Legal Analysis* 1–45.

Smith, Jeremiah. 1911. Legal Cause in Actions of Tort. 25 *Harvard Law Review* 103-128.

Wright, Richard W. 1985. Actual Causation vs. Probabilistic Linkage: The Bane of Economic Analysis. 14 *Journal of Legal Studies* 435-456.

Wright, Richard W., and Ingeborg Puppe. 2016. Causation: Linguistic, Philosophical, Legal and Economic. 91 *Chicago-Kent Law Review* 461-502.

Appendix

Cause-in-Fact Requirements in a Model of Liability for Negligence: Formal Analysis

A.1 Assumptions and framework of analysis

A party chooses among three actions: not to engage in an activity; to engage in the activity and to exercise care; or to engage in the activity and not to exercise care.

If the party does not engage in the activity, he obtains no payoff.

If he does engage in the activity, he obtains a certain payoff of $v > 0$. Also, if he takes care when participating in the activity, he bears a cost of $c > 0$. Hence, if he engages in the activity and takes care, his payoff is $v - c$.

Three states of the world may occur: a safe state, S, occurring with probability p_S; a risky state R, occurring with probability p_R; and a dangerous state, D, occurring with probability p_D. The probabilities of the states are assumed to be positive and to sum to 1.

Two possible outcomes can occur: either no harm will transpire or harm of a fixed positive magnitude h will arise. Which outcome eventuates depends on the party's action and the state of the world as displayed in Table 1.

A.2 Socially desirable behavior

Social welfare is defined to be the expected value of participation in the activity minus the cost of care and of harm that might occur. We can now determine social welfare for each of the three possible actions and determine which is best given p_S, p_R, p_D, v, c, and h.

If the action is not to engage in the activity, social welfare will be 0. If the action is to engage in the activity and to exercise care, social welfare will be

(1) $v - p_D h - c$

because harm will occur only in state D. If the action is to engage in the activity but not to take care, social welfare will be

(2) $v - (p_R + p_D)h$

because harm will occur in states R and D. We therefore have the following.

Proposition 1. (a) Not participating in the activity is socially best when both

(3) $v < p_D h + c$

and

(4) $v < (p_R + p_D)h$

hold.

(b) Participating in the activity and taking care is socially best when

(5) $v > p_D h + c$

and

(6) $c < p_R h$

hold.

(c) Participating in the activity and not taking care is socially best when

(7) $v > (p_R + p_D)h$

and

(8) $c > p_R h$

hold.

Comment. For simplicity, I am not stating here (or below) conditions under which different actions might result in the same level of social welfare.

Proof. The claims are self-evident. For example, not participating is best when it is superior to participating and exercising care (meaning that (1) is negative) and superior to participating and not exercising care (meaning that (2) is negative). □

A.3 Behavior under different cause-in-fact requirements for liability under the negligence rule—when law enforcement is perfect

For a party who is engaging in an activity, let us define *negligence as failure to exercise care when doing so would raise social welfare*. If a party who engages in the activity takes care, social welfare will be (1), whereas if he does not take care social welfare will be (2). It follows that if a party fails to take care when $c < p_R h$, social welfare would be increased if care were exercised; thus failure to take care when $c < p_R h$ means that the party is negligent.

Suppose that a party was negligent and that harm occurred. We will say that his *negligence was a cause-in-fact of the harm if the harm would not have occurred had he exercised care*. It is clear that this will be the case only in state *R*.

We define *the negligence rule with the negligence-based cause-in-fact requirement to be the regime under which a negligent party would be found liable and would thus have to pay damages of h when harm occurred if and only if his negligence was a cause-in-fact of harm*. In the present section we will assume that law enforcement of this regime is perfect in the sense that liability will always be found when it is authorized. We now have

Proposition 2. Suppose that the negligence rule with the negligence-based cause-in-fact requirement applies. Then if the party engages in his activity, he will be induced to take care when doing so is socially desirable but not would not take care otherwise.

Proof. If taking care is socially desirable and the party takes care, he will not be negligent so will not be found liable if harm occurs. Hence his utility will be

(9) $v - c$.

If he does not take care, he will be negligent and his negligence will be found a cause-in-fact of harm in state *R*. Hence his expected utility will be

(10) $v - p_R h$.

But since taking care is socially desirable, $c < p_R h$. Hence,

(11) $v - p_R h < v - c$.

Accordingly, the party would exercise care.

If taking is not socially desirable, however, the party would never be found negligent and thus would not exercise care. (He will engage in the activity and obtain utility of *v*.) □

Next, and supposing again that a party who engaged in an activity was negligent and that harm occurred, let us say that the party's *activity was a cause-in-fact of harm if the harm would not have occurred had he not participated in his activity*. This will be the case in states *R* and *D*. We define *the negligence rule with the activity-based cause-in-fact requirement to be the regime under which a negligent party would be found liable and would thus have to pay damages of h when harm occurred if and only if his activity was a cause-in-fact of harm*. We then have

Proposition 3. Suppose that the negligence rule with the activity-based cause-in-fact requirement applies. Then if the party engages in his activity, he will be induced to take care when doing so is socially desirable but would not take care otherwise.

Proof. If taking care is socially desirable and the party takes care, he will not be negligent, so that his utility will be given by (9). If he does not take care, he will be negligent and his expected utility will be given by (2). But

(12) $v - (p_R + p_D)h < v - c$

because $c < p_R h$ holds. Thus, the party would exercise care.

If taking is not socially desirable, the party would never be found negligent and thus would not exercise care. □

Last, we consider the issue of participation in the activity.

Proposition 4. Suppose that the negligence rule applies with either the negligence-based cause-in-fact requirement or the activity-based cause-in-fact requirement. Then

(a) a party's decision whether to engage in the activity will be the same under both requirements;

(b) a party will always decide to engage in the activity when that would be socially desirable; but

(c) a party will sometimes decide to engage in the activity when that would be socially undesirable: if the exercise of care would be socially desirable, a party will inappropriately engage in the activity when *v* is in $(c, c + p_D h)$; and if the exercise of care would be socially undesirable, the party will inappropriately engage in the activity when $v < (p_D + p_M)h$.

Proof. (a) The decision whether to engage in the activity depends on a party's utility if he does not engage in the activity and his utility if he engages in the activity. These utilities do not

depend on which cause-in-fact requirement applies. In particular, if a party does not engage in the activity, his utility will be 0 regardless of the cause in fact requirement. And if a party does engage in the activity, Propositions 2 and 3 tell us that, regardless of the requirement, he will take care when that is socially desirable—and thus obtain $v - c$—and not take care when that is socially undesirable—and thus obtain v. The party will therefore behave identically under both requirements.

(b) When the activity is socially desirable, one case is where care is also socially desirable. From Propositions 2 and 3 we know that in this case, regardless of the cause-in-fact requirement, if the party engages in the activity, he will take care and not bear liability. Thus he will engage in the activity if $v > c$. And we know from Proposition 1 that since the activity is socially desirable, $v > c + p_D h$, implying that $v > c$ so that the party will engage in the activity.

The other case is where care is not socially desirable. From Propositions 2 and 3 we know that in this case, regardless of the requirement, if the party engages in the activity, he will not take care and not bear liability. Thus he will engage in the activity if $v > 0$, which is true by assumption.

(c) We again consider two cases. One is where taking care is socially desirable. Then engaging in the activity is socially undesirable if $v < c + p_D h$. But in this case the party will bear no liability if he takes care; and since he has the option to do this, he will engage in the activity whenever $v > c$. Hence, we will observe inappropriate engagement in the activity when v is in (c, $c + p_D h$) under both cause-in-fact requirements.

The second case is where taking care is socially undesirable. In this case, engaging in the activity is socially undesirable when $v < (p_D + p_M)h$. But because there would be no liability in this case, the party will engage in the activity so long as v is positive. Hence, regardless of the cause-in-fact requirement there will be engagement in the activity that is socially undesirable when $v < (p_D + p_M)h$. □

A.4 Behavior under different cause-in-fact requirements for liability under the negligence rule—when law enforcement is imperfect

We now reconsider behavior given the two cause-in-fact requirements under the assumption that law enforcement is imperfect in the sense that a party who ought to be found liable and pay damages will be successfully sued and pay damages only with a probability q—and thus that with probability $1 - q$ a party will escape liability. The next two propositions concern the exercise of care under the cause-in-fact requirements given q.

Proposition 5. Suppose that the negligence rule with the negligence-based cause-in-fact requirement applies with probability q, that the exercise of care is socially desirable, and that a party engages in his activity. Then the party will be induced to take care only when q exceeds a critical value $q_N = c/(p_M h)$; when q equals q_N, the party will be indifferent between taking care and not; and when q is less than q_N, he will not take care.

Proof. If taking care is socially desirable and the party takes care, he will not be negligent so will not be found liable if harm occurs. Thus, his utility will be (9). As observed in the proof

of Proposition 2, we know that if a party engages in the activity and is negligent, his expected liability will be (10), meaning that when law enforcement is imperfect, his expected liability will be $v - qp_Mh$. Thus the expected utility of a party who engages in the activity and takes care minus his expected utility if he does not is

(13) $v - c - (v - qp_Mh)$.

Accordingly, if $q = c/(p_Mh)$, (13) will be zero and a party who engages in the activity will be indifferent between exercising care and not; if $q < c/(p_Mh)$ the party will not take care; and if $q > c/(p_Mh)$, the party will take care. Note that $c/(p_Mh) < 1$ because care is (8) is assumed to hold.

If taking care is not socially desirable, the party would not be found negligent and would thus not take care. □

Proposition 6. Suppose that the negligence rule with the activity-based cause-in-fact requirement applies with probability q, that the exercise of care is socially desirable, and that a party engages in his activity. Then the party will be induced to take care only when q exceeds a critical value $q_A = c/[(p_D + p_M)h]$; when q equals q_A, the party will be indifferent between taking care and not; and when q is less than q_A, he will not take care.

Proof. If taking care is socially desirable and the party takes care, he will not be negligent so will not be found liable if harm occurs. Thus, his utility will be $v - c$. As observed in the proof of Proposition 3, we know that if a party engages in the activity and is negligent, his expected liability will be (1), meaning that when law enforcement is imperfect, his expected liability will be $v - q(p_D + p_M)h$. Hence, the expected utility of a party who engages in the activity and takes care minus his expected utility if he does not is

(14) $v - c - (v - q(p_D + p_M)h$.

Therefore, if $q = c/[(p_D + p_M)h]$, (14) will be zero and a party who engages in the activity will be indifferent between exercising care and not; if $q < c/[(p_D + p_M)h]$, the party will not take care; and if $q > c/[(p_D + p_M)h]$, the party will take care. Furthermore, $c/[(p_D + p_M)h] < 1$ because (8) is assumed to hold.

If taking care is not socially desirable, the party would not be found negligent and would thus not take care. □

From the previous two propositions, we see that $q_A < q_N$. This leads to the next conclusion.

Proposition 7. Suppose that the negligence rule applies with probability q, that the exercise of care is socially desirable, and that a party engages in his activity. Then the party will be induced to exercise care more often under the activity-based cause-in-fact requirement than under the negligence-based cause-in-fact requirement. In particular, the party will be negligent under the negligence-based cause-in-fact requirement but not under the activity-based cause-in-fact requirement whenever q is in the interval (q_A, q_N).

Finally, let us consider engagement in the activity under the two cause-in-fact requirements.

Proposition 8. Suppose that the negligence rule with either of the cause-in-fact requirements applies with probability q. Then

(a) a party will always engage in his activity when doing so is socially desirable;

(b) a party will sometimes engage in his activity when doing so is socially undesirable;

(c) a party will undesirably engage in his activity at least as often under the negligence-based cause-in-fact requirement as under the activity-based cause in fact requirement—if the party undesirably engages in his activity under the activity-based cause-in-fact requirement, he will also do so under the negligence-based cause-in-fact requirement; and

(d) there exists a range of situations in which the exercise of care is socially desirable and under the negligence-based cause-in-fact requirement, a party will socially undesirably engage in the activity and fail to take care, whereas under the activity-based cause-in-fact requirement, he will socially desirably engage in the activity and take care.

Comment: Parts (c) and (d) imply that parties will undesirably engage in the activity more often under the negligence-based cause-in-fact requirement than under the activity-based cause-in-fact requirment.

Proof. (a) It was shown in Proposition 4(b) that a party would always engage in his activity when $q = 1$, meaning that his achievable utility is positive. When $q < 1$, the maximum utility of a party who engages in the activity must be at least as high as it was when $q = 1$, and thus must be positive. Therefore, he will engage in the activity.

(b) The proof of Proposition 4(c) applies when $q < 1$.

(c) Suppose that under activity-based cause-in-fact requirement, the party undesirably engages in the activity. There are then two possibilities—the party exercised care; or the party did not exercise care. It will be shown that under each possibility, the party would undesirably engage in the activity under the negligence-based cause-in-fact requirement.

(i) The party exercised care: Because the hypothesis is that under the activity-based cause-in-fact requirement engaging in the activity was socially undesirable and care was exercised, we know in particular that $v < c + p_D h$.

Because the party chose to take care, we know that doing so must have been socially desirable, that is, that $c < p_M h$. The logic is that if taking care was not socially desirable, the party could not be found negligent for failure to take care. Hence he would not have taken care, a contradiction.

Because the party exercised care and thus could not have been found liable, and because he engaged in the activity, we know that $v > c$.

Now let us use the above inferences to show that under the negligence cause-in-fact requirement, the party would also have undesirably engaged in the activity.

First, since $v > c$, we know that under the negligence-based cause-in-fact requirement, the party would choose to engage in the activity: The party has the option to exercise care and avoid liability. Thus, if he engages in the activity and takes care, his return will be better off than not not engaging in the activity, for $v > c$. (Although we cannot infer from what was just said that he will choose to take care, we do know that since he has the *option* to do so, if he acts optimally it must involve engaging in the activity.)

Second, let us show that the party's engagement must be socially undesirable. One possibility is that the party engages in the activity and takes care. But since we know $v < c + p_D h$, doing so would be socially undesirable. The other possibility is that the party engages in the activity and does not take care. In this case, engagement in the activity will be undesirable if $v < p_M h + p_D h$. But this condition must hold because we showed that $v < c + p_D h$ and $c < p_M h$.

(ii) The party did not exercise care: Because the hypothesis is that under the activity-based cause-in-fact requirement engaging in the activity was socially undesirable and care was not exercised, we know in particular that $v < (p_M + p_D)h$.

The inference we draw from the hypothesis that the party did not take care depends on whether care was not or was socially desirable. If care was not socially desirable, there would be no liability for failing to take care. Hence, the party would engage in the activity as long as v is positive. If care was desirable, however, the party would be negligent and thus face expected liability of $q(p_M + p_D)h$ under the activity cause-in-fact requirement. And because the party did not take care, we infer that $c > q(p_M + p_D)h$. Finally, because the party chose to engage in the activity, we infer that $v > q(p_M + p_D)h$.

Now we can verify that under the negligence-based cause-in-fact requirement, the party will also undesirably engage in the activity. Suppose first that care is not socially desirable. In that case, the party will have no reason to take care and thus will engage when $v > 0$. Thus the party will engage in the activity and that will be socially inappropriate given our assumption that $v < p_M h + p_D h$.

Next suppose that care is socially desirable. We had inferred above that $c > q(p_M + p_D)h$. This implies that $c > qp_M h$, meaning that under the negligence-based causal requirement, the party will choose not to take care. And we had also inferred that $v > q(p_M + p_D)h$. This implies that $v > qp_M h$, meaning that the party will engage in the activity under the negligence-based requirement. And since our hypothesis is that $v < (p_M + p_D)h$, engaging in the activity will be socially undesirable.

(d) Suppose that the exercise of care is socially desirable given engagement in the activity, that is, suppose that $c < p_M h$.

We know from Propositions 5 and 6 that if q lies in the interval $(q_A, q_N) = (c/[(p_D + p_M)h], c/(p_M h))$, then the party will exercise care under the activity-based cause-in-fact requirement but will not do so under the negligence-based cause-in-fact requirement.

Now consider any q in the interval just given and any v in the interval $(c + p_D h, (p_M + p_D)h)$. We claim that for such q and v, a party will undesirably engage in the activity and fail to take care under the negligence-based requirement, whereas he will desirably engage in the activity and take care under the activity-based requirement. The claims about the taking of care are true, as explained in the prior paragraph.

Regarding the claim that the party will undesirably engage in the activity under the negligence-based cause-in-fact requirement, we know that if he engages in the activity, he will be negligent, and thus his expected liability will be $qp_M h$. But since $q < c/(p_M h)$, we have $qp_M h <$

c. Thus, expected liability is less than *c.* But $v > c$ in the interval. Hence the party will engage in the activity. That will be socially undesirable because in the interval $v < (p_M + p_D)h$.

Finally, regarding the claim that the party will desirably engage in the activity under the activity causal requirement, we know that if he engages in the activity, he will take care and thus his expense will be *c*. But, again, $v > c$ in the interval, so the party will engage in the activity. Further, that will be socially desirable since $v > c + p_D h$ in the interval. □