

# Risk Aversion and the Optimality of Attenuated Legal Change

Steven Shavell\*

*The central point made here is that legal change should often be attenuated — should be less rigorous than conventionally efficient legal change — in order to reduce risk-bearing burdens. This conclusion rests on two arguments. First, insurance against legal change is largely unavailable (due primarily to the correlated nature of the losses usually caused by legal change). Second, given the unavailability of such insurance coverage, it is desirable in principle for legal change to be less than conventionally efficient when the parties governed by it are risk-averse.*

## 1. Introduction

One of the major risks that individuals and firms face in the modern world is of legal change. Homeowners, for example, confront the prospect that the mortgage interest tax deduction will be eliminated; manufacturers bear the risk that regulations concerning workplace safety or environmental harm will be modified; and fishermen live with the possibility that a fishery will be declared off-limits. The risk of significant legal change seems inevitable, as it is a natural byproduct of uncertain economic and technological change and of the emergence of new information.

A hallmark of the risk of legal change is that it is largely uninsurable. As is discussed in Section 2, insurance against legal change is for the most part nonexistent (whereas insurance against other substantial risks is usually widespread). This fact is suggested to be attributable to an important degree to the correlated nature of the losses associated with legal change — all homeowners who now deduct mortgage interest from their taxable income would simultaneously suffer financial harm if their present tax benefit were abolished, which could impose an intolerable risk on insurers.

Given this background, it is assumed that insurance against legal change is unavailable in a model that is analyzed in Section 3. Under that assumption as well as the assumption that parties subject to the law are risk averse, the basic point of the article is developed: that legal change should be attenuated.

To amplify, in the model, there is a risk that an activity will be discovered to be harmful by the state, in which event a law addressing the harmful activity will be adopted. Individuals are able to reduce the harmfulness of the activity by exercising costly care.

---

\* Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School. I thank A. Mitchell Polinsky for comments on this article, Kenneth Abraham, Tom Baker, David Bassi, Patricia Danzon, Scott Harrington, and Micah Woolstenhulme for discussions about insurance, Edward Dumoulin, David Geiger, and Jodie-Jane Tingle for research assistance, and the John M. Olin Center for Law, Economics, and Business at Harvard Law School for research support.

One form of law that is considered is regulation, under which the state mandates a level of care. Regulation may be interpreted either as direct control of precautions by a government agency or as effective control by the courts through the application of the negligence rule. The uninsurable risk borne by individuals involved in the activity is that, if the activity is revealed to be harmful, regulation will be adopted and they will have to spend to meet the required level of care. It is shown that the optimal regulatory standard of care is less than the conventionally efficient level, that is, less than that following from the standard cost-benefit calculus.

The rationale for this result is, in essence, that if the stipulated standard of care were the conventionally efficient level, then a marginal relaxation of the standard would leave expected social costs essentially unchanged,<sup>1</sup> but the reduction of the standard would produce a social benefit by lowering risk-bearing for the risk-averse parties subject to the standard. Thus, it is always socially desirable for the standard of care to be less than (and perhaps to be much less than) the level that would be conventionally efficient.

The other form of law that is analyzed is a required payment based on the harm caused or estimated. The interpretation of this form of law is the payment of damages to victims under the rule of strict liability, or the payment of a fine or a corrective tax to the state. Here the risk that individuals subject to the law face is both having to make payments and having to bear a cost of care (for individuals will be induced to take care if the activity is found to be harmful). It is shown that the optimal magnitude of the payment is less than the harm (or the expected harm).<sup>2</sup> The rationale for this conclusion is similar to that for the conclusion about attenuation of the optimal standard of care.

In Section 4, several concluding comments about the model are made, and it is noted that the theoretical results about attenuation exhibit some consistency with reality in that observed legal change is often attenuated through partial or delayed implementation and through grandfathering.

This article was stimulated by an economically-oriented, mainly informal literature on legal change, beginning with articles by Graetz (1977) and Kaplow (1986), who emphasize the view that legal change should not be attenuated, and continuing with Levmore (1999), Logue (2003), Masur and Nash (2010), and Shavero (2000), among others.<sup>3</sup> The main reason for the difference between the conclusions that I draw and those of much of the prior literature is that

---

<sup>1</sup> At the conventionally efficient level of care, the marginal gain from a reduction in the level of care must equal the marginal loss from the resulting increase in expected harm.

<sup>2</sup> This is true whether victims of harm are risk neutral or risk averse, as shown in Propositions 3 and 4. It is also demonstrated that the optimal level of payment is zero for all harms sufficiently low; see Proposition 3.

<sup>3</sup> See also Kaplow (1992), Levmore (1998), Logue (1996), Nash and Revesz (2007), Quinn and Trebilcock (1982), and Shavell (2008). Some of the literature on legal change does provide rationales for attenuated legal change — for example, Levmore (1999), Logue (2003), and Shavell (2008) — but these rationales are different from the absence of insurance coverage against legal change.

past work overlooks, or does not adequately recognize, the lack of availability of insurance coverage against legal change.<sup>4</sup>

## 2. The Absence of Insurance Coverage against Legal Change

Insurance against legal change is not generally offered by the insurance market. That this is so can be verified in a number of ways. First, the categories of coverage that insurers state that they sell do not include legal change. For example, insurers such as Allstate, Geico, and Travelers list as major categories of coverage property, life, health, auto, homeowners, liability, and a number of others, but none mention legal change as a type of coverage.<sup>5</sup> Second, reference works and texts on insurance do not describe legal change as a distinct risk for which coverage is available.<sup>6</sup> Third, knowledgeable individuals whom I have contacted in the insurance industry and in academia concur that insurance coverage against legal change is essentially unavailable.<sup>7</sup>

Thus, if we ask whether the risks due to legal change that I mentioned at the outset are insurable, the answer is basically no. It does not seem possible for homeowners to purchase insurance coverage against elimination of the mortgage interest tax deduction, for manufacturers to secure coverage against the cost of meeting new workplace safety or pollution regulations, or for fishermen to obtain coverage against a ban on fishing.<sup>8</sup>

---

<sup>4</sup> With one exception, none of the references on legal change cited above discuss the fact that insurance coverage against legal change is unavailable. The exception is Masur and Nash (2010), but they do not regard the lack of coverage as an important basis for relief from legal change, which they state should be “quite rare” at p. 402. In a related vein, Blume and Rubinfeld (1984), writing on regulatory takings of land, observe that insurance coverage against these takings does not exist, but do not argue that some form of relief is broadly justified as a consequence.

<sup>5</sup> For an overview of Allstate’s insurance products, see [www.allstate.com/about/product-overview.aspx](http://www.allstate.com/about/product-overview.aspx); for Geico, see [www.geico.com](http://www.geico.com) and the pull-down menu “type of insurance”; for Travelers, see [www.travelers.com](http://www.travelers.com).

<sup>6</sup> I have examined Appleman (1996) and Couch (2009), major treatises on insurance law; Abraham (2010) and Baker (2008), leading casebooks on insurance law and policy; and Harrington and Niehaus (2004) and Vaughan and Vaughan (2008), well-known textbooks on the business and economics of insurance. These resources cover mainly the following major categories of insurance: health, disability, life, homeowners, marine, commercial, automobile, fire, property, theft, and liability. None discusses legal change as a separate area of coverage.

<sup>7</sup> These individuals include David Bassi of Plymouth Rock Assurance Company and Micah Woolstenhulme of Guy Carpenter & Company, Kenneth Abraham and Tom Baker, legal academics whose focus is on insurance law, and Patricia Danzon and Scott Harrington, academic economists whose primary research is on insurance. (They also agree with the essence of the qualification that I make below about the coverage against legal change that is bound up in liability insurance policies.)

<sup>8</sup> Although to my knowledge no insurance exists against a long term ban on fishing, it might be asked whether business interruption coverage would compensate fishermen against a temporary ban, such as for a period after an oil spill due to the risk of contamination. The answer appears to be negative, mainly because business interruption coverage ordinarily requires that a loss be associated with property owned by the insured and that this property have actually sustained physical damage. See, for example, Abraham (2010) at 226-232 and Baker (2008) at 41 and 314-321.

There is, however, a partial exception to the absence of coverage against legal change worthy of note. Namely, significant protection against modifications in liability rules is implicit in standard liability insurance policies.<sup>9</sup> Although the primary role of liability insurance is to cover insureds against the risk of liability resulting from existing laws, liability insurance also covers insureds against changes in liability rules<sup>10</sup> as long as the changes concern the insured categories of liability. Hence, physicians and other professionals would likely be covered under their malpractice policies against expansions in their exposure to liability. Nevertheless, this type of protection against changes in liability rules is incomplete: it does not compensate insureds for the costs of any additional precautions that they are led to take, and it does not prevent insurers from raising premiums,<sup>11</sup> from excluding new liability risks at the time of policy renewal,<sup>12</sup> or from canceling coverage altogether.<sup>13</sup>

Another qualification to the unavailability of insurance coverage against legal change is that parties can sometimes hedge against it. For instance, homeowners facing the risk of elimination of the mortgage interest deduction might consider selling short shares in home building companies, for these companies would be expected to suffer from a fall in demand for new homes if the deduction was disallowed. Yet such hedging opportunities are often limited<sup>14</sup> and, in any case, require a degree of sophistication that many parties do not possess.

What is the explanation for the fact that insurance coverage against legal change is largely unavailable? Are there distinctive aspects of legal change that set it apart from the broad swath of risks — from automobile accidents, to fires, to theft — for which insurance coverage is widely sold?

A salient characteristic of legal change is that it often affects all individuals subject to a law at the same time. Thus, as I observed, all individuals who now deduct mortgage interest

---

<sup>9</sup> There also may be occasional instances of sale of coverage against explicit risks of legal change. For instance, David Bassi mentioned to me that coverage against specific changes in the tax laws have been sold on a negotiated basis to certain taxpayers. Such examples seem to be of negligible importance.

<sup>10</sup> Baker (2004) makes this point in a broad discussion of the risks of change that liability insurance covers.

<sup>11</sup> See, for example, Priest (1987), finding that insurers raised liability insurance premiums dramatically due to the expansion of tort liability, Viscusi (1991) at 176-177, stating that insurers raised product liability insurance premiums due to heightened product liability exposure, and Abraham (1977) at 489-491, suggesting that insurers increased medical malpractice premiums due to increased risks of liability for medical malpractice.

<sup>12</sup> See, for example, Abraham (2010) at 570-575 and Harrington and Niehaus (2004) at 616-617, noting the greater use of exclusions for pollution-related liability due to a broadening of the legal grounds for such liability.

<sup>13</sup> See, for example, Church (2005), discussing the cancellation of a wide array of types of liability insurance policies on account of greater liability risks.

<sup>14</sup> For instance, the ability to sell short shares in home building companies is circumscribed because the aggregate capitalization of these companies is only a fraction of the total risk facing owners of mortgages from elimination of the deduction, because there are costs of selling short, and because selling short requires an individual to have a brokerage account with short selling privileges. Moreover, selling short involves its own risks — the price of shares in home building companies could rise.

from their taxable incomes would together suffer a loss were this tax benefit eliminated. Hence, an insurer selling coverage against that legal change would bear a large risk. Suppose that an insurer writes \$1 billion of coverage against elimination of the mortgage interest tax deduction. Then the insurer must maintain reserves fully equal to \$1 billion in order to honor its contracts with policyholders — regardless of how low the probability of claims is, such as only 1%, the triggering of its coverage responsibilities would occur simultaneously for all insureds.<sup>15</sup> In contrast, if an insurer writes \$1 billion worth of coverage against, say, car theft involving a 1% risk of a \$25,000 loss per car that is independent across car owners, the insurer could maintain reserves of only \$12 million and be able to honor its contracts with virtual certainty.<sup>16</sup> This pronounced difference in required reserves illustrates why positive correlation of losses is a stock reason given in the insurance literature for the nonexistence of coverage.<sup>17</sup>

Correlation of losses therefore seems to provide a credible explanation for the unavailability of coverage against a risk like that of elimination of the tax deductibility of mortgage interest. The magnitude of the financial threat that such a risk would pose for insurers would probably be too great for insurers to bear,<sup>18</sup> and the insurance experts to whom I have spoken agree broadly with this belief. Still, correlation of losses needs to be supplemented with other factors to help explain the absence of coverage for the many legal risks that are not of the scope of the elimination of the mortgage interest deduction.<sup>19</sup>

---

<sup>15</sup> An insurer might choose to hold such reserves in order to prevent bankruptcy (and to give insureds confidence in their insurance), or the insurer might be forced to hold the reserves by a regulator.

<sup>16</sup> Since the assumption is that the insurer writes coverage of \$1 billion and each individual's loss from car theft would be \$25,000, the insurer must be covering 40,000 individuals. The standard deviation of an individual's loss if not insured is \$2,487 (namely,  $\sqrt{.01(\$25,000 - \$250)^2 + .99(-\$250)^2}$ ), implying that the standard deviation of the average loss among the 40,000 insured individuals is only  $\$2,487/\sqrt{40,000} = \$12.44$ . Now if the insurer has \$12 million in reserves, it will have reserves of \$300 per individual. Hence, for the reserves to be exhausted, the average loss must exceed \$300. Since the mean loss per individual is \$250, the average loss would have to exceed the mean by at least \$50 for reserves to be depleted. But  $\$50/\$12.44$  is 4.02, which is to say, more than four standard deviations. The odds of such an event are less than one in 10,000 for a normal distribution.

<sup>17</sup> See, for example, Harrington and Niehaus (2004) at 182, and Vaughan and Vaughan (2008) at 43. It should also be noted that positive correlation of losses can lead to uninsurability whether or not one views insurers as risk neutral or risk averse. Because correlation of losses implies that required reserves are high, insurers' capital costs become high. Hence, even if insurers are taken to be risk neutral, they must charge a premium substantially above the actuarially fair premium to cover their capital costs. That would reduce demand for coverage to a level below full coverage, and possibly to zero. If insurers are regarded as risk averse, they would charge greater premiums, further suppressing demand.

<sup>18</sup> The total value of the deductions is estimated to be from \$70 billion to \$100 billion in a single year. See, for example, Wallace (2011). Since individuals would presumably often want to insure against the value of their losses over the lifetime of their mortgages, insurers could face the risk of total claims of hundreds of billions of dollars. This could make the exposure of the insurance industry greater than that from any single event in its history.

<sup>19</sup> Consider, for instance, a change in the tax laws that would apply to only a small subset of farmers or a change in a municipality's fire code that would pertain only to local fast food restaurants. These legal risks do not seem too large for the insurance industry to assume, and to some extent they could be set off against each other (a change in a tax law affecting farmers might be uncorrelated with a change in a fire code applying to fast food

The view that legal change usually leads to correlation of losses and accordingly to the absence of insurance coverage is, it should be made clear, reconcilable with the fact that liability insurance policies provide a kind of coverage against changes in liability rules. The reason is that a change in a liability rule (unlike a change in a tax or a regulatory requirement) does not lead to strong correlation of losses: a change in a liability rule does not imply losses for all parties subject to the rule. Indeed, when a liability rule changes, typically only a small minority of insureds will make claims as a consequence, for only those who turn out to cause harm and are found liable as a result of the modification of the rule would add to the claims the insurer must pay.<sup>20</sup>

A further point suggests that correlation of losses is the primary explanation for the general absence of insurance coverage against legal change: neither of the other two standard reasons for the nonexistence of insurance coverage — moral hazard and adverse selection — stand out as important. Moral hazard should not be significant because the sale of coverage to an individual would not be likely to alter the likelihood of a change in the law;<sup>21</sup> and adverse selection should not be of weight because individuals all face the same risk of legal change and, in any event, would not be likely to know more about the chance of legal change than insurers.

To summarize and conclude the discussion in this section, we have seen that insurance coverage against legal change is largely unavailable (with the proviso concerning liability insurance), and we have a reasonably satisfying primary explanation for this fact, the correlation of losses. I will therefore assume in the model below that insurance against legal change does not exist.<sup>22</sup>

### 3. The Model

Parties called injurers engage in an activity that may turn out to be harmful to parties called victims. In particular, the state learns at a future date whether the activity of injurers is harmful. If it is harmful, the state announces a legal rule to address the danger. After the

---

restaurants). Yet such legal risks are not insurable. Part of the explanation may be that insurers (and reinsurers) would still need reserves equal to the full amount of the coverage written to assure that they could satisfy their contracts, making provision of the coverage expensive. See also Froot (2001), addressing the closely related question of the explanation for the paucity of insurance coverage of catastrophic risks, such as earthquakes, floods, and hurricanes (which, like legal changes, also affect many parties at once).

<sup>20</sup> Furthermore, as I mentioned above, liability insurers have the escape hatch of raising premiums, excluding new exposures to liability, or discontinuing coverage at the time of policy renewal.

<sup>21</sup> A different observation is that if a mass of individuals were covered against a legal change, legislators or courts might be more likely to effect the change, since it would impose less hardship on those subject to it and since they would be less likely to resist it. But such a general tendency would not imply that the purchase of insurance by any single individual would be discouraged, for the tendency would not imply that the individual would raise the risk of legal change (and thus his premium) by his own, personal purchase of coverage.

<sup>22</sup> However, when a change in a liability rule occurs, I do allow for individuals to purchase coverage against liability under the new rule for the reasons discussed above. Nevertheless, individuals still bear risk: the risk of having to pay the premium for the new coverage and of having to spend on care under the new rule.

announcement of the rule, injurers choose a level of care to reduce the probability of harm. Injurers are identical to each other, as are victims.

Specifically, define

$q$  = probability that the state learns that the activity is harmful and announces a legal rule;

$$0 < q < 1;$$

$x$  = expenditure on care by an injurer to reduce the probability of an accident if the activity is harmful;  $x \geq 0$ ;

$p(x)$  = probability of an accident if the activity is harmful;  $0 < p(x) < 1$ ;  $p'(x) < 0$ ;

$$p''(x) > 0; \text{ and}$$

$h$  = harm if an accident occurs;  $h > 0$ .

If the activity is harmful, the sum of care and expected harm is

$$(1) \quad x + p(x)h.$$

Let the  $x$  that minimizes (1) be denoted  $x^*$ ; and call  $x^*$  the conventionally optimal level of care because in models with risk neutral actors, it is usually assumed that minimization of (1) is the social objective.<sup>23</sup> Because the case in which  $x^*$  is zero is uninteresting, it will be supposed that  $x^*$  is positive, or equivalently, that

$$(2) \quad p'(x)h = -1$$

holds for a positive  $x$ . Note that since  $p''(x) > 0$  for  $x \geq 0$ , condition (2) holds for a positive  $x$  if and only if

$$(3) \quad p'(0)h < -1,$$

so that this inequality will be assumed.

Let

$U(\cdot)$  = utility of wealth of an injurer, and

$u$  = initial wealth of an injurer.

Injurers are assumed to be risk averse or risk neutral, but the case in which they are risk averse will be emphasized because the risk aversion of those subject to the law is the main concern of this article. When injurers are assumed to be risk neutral, the utility of an injurer will be taken to equal the amount of his wealth.

It is assumed that injurers cannot purchase insurance against legal change for the reasons given in section 2. The meaning of this assumption will be discussed further below.

Also, let

$V(\cdot)$  = utility of wealth of a victim, and

$v$  = initial wealth of a victim.

Both the case in which victims are risk neutral (with the utility function of a victim then taken to equal the amount of his wealth) and the case in which they are risk averse will be analyzed. The case of risk neutral victims is considered both because it is expositionally simpler (the effect of injurer risk aversion on optimal legal change is most easily understood in isolation from the possible risk aversion of victims) and because it is sometimes descriptive of reality (when the harm to each victim is limited or, often, when the government bears the harm). The qualitative

---

<sup>23</sup> See, for example, Landes and Posner (1987) and Shavell (1987).

nature of the results does not change, however, when victims are risk averse. Assumptions about victims' ability to insure in the case in which they are risk averse will also be discussed below.

We will determine Pareto optimal legal rules. A rule is defined to be Pareto optimal for the parties given an initial situation if there does not exist an alternative rule, and a transfer payment between injurers and victims, under which the expected utility of both injurers and victims would be higher. To identify Pareto optimal legal rules, it is necessary and sufficient to solve the following problem:

- (4) Maximize the expected utility of injurers  $EU$  over possible versions of a legal rule, subject to the constraint that
- (5) the expected utility of victims  $EV$  is held constant by means of a transfer payment by injurers to victims,

where

$t$  = transfer payment by injurers to victims.

It is assumed that  $t$  is made before the state learns whether the activity is harmful, that is, before legal uncertainty is resolved.

Of course, it does not make sense to treat minimization of  $x + p(x)h$  as the optimality criterion, for that objective does not reflect risk-bearing by risk-averse parties. (As will be observed, however, the problem of maximizing (4) subject to (5) reduces to minimization of  $x + p(x)h$  when both injurers and victims are risk neutral.)

We now consider Pareto optimal legal rules when the rules concern regulation of behavior and when they concern payments for harm.

### 3.1 Regulation of Behavior

Assume here that if the state learns that injurers' activity is harmful, the legal rule that the state adopts is regulation, by which is meant a standard of care that injurers must exercise. Let

$x_s$  = required standard of care if the state learns that injurers' activity is harmful.

It will be assumed that injurers meet the standard  $x_s$ , for consideration of its method of enforcement by the state would be distracting for our purposes. (As noted in the introduction, enforcement could occur through direct control by a regulatory agency or through the use of the negligence rule under the liability system.)

Consider first the determination of a Pareto optimal standard of care assuming that victims are risk neutral. The expected utility of an injurer is then

$$(6) \quad EU = (1 - q)U(u - t) + qU(u - t - x_s).$$

Note that this expression reflects the assumption that the transfer payment  $t$  is made before the state learns whether the activity is harmful, that the risk borne by an injurer is having to spend  $x_s$  to meet the standard, and that injurers do not have insurance coverage against the risk of  $x_s$ . The expected wealth of a victim is  $v + t - qp(x_s)h$  and must satisfy

$$(7) \quad v + t - qp(x_s)h = r,$$

where  $r$  is a reference level of expected wealth. The Pareto optimal standard  $x_s$  maximizes (6) subject to (7).<sup>24</sup> From (7), we have

$$(8) \quad t = r - v + qp(x_s)h,$$

so the problem at issue reduces to maximizing

$$(9) \quad EU(x_s) = (1 - q)U(u - r + v - qp(x_s)h) + qU(u - r + v - qp(x_s)h - x_s)$$

over  $x_s$ . Denote the Pareto optimal standard  $x_s$  that maximizes (9) by  $x_s^*$ .<sup>25</sup>

Before determining the Pareto optimal standard when injurers are risk averse, let me note what the Pareto optimal standard is when injurers are risk neutral.

*Remark 1.* Assume that both injurers and victims are risk neutral. Then the Pareto optimal standard of care  $x_s^*$  is the conventionally optimal level of care  $x^*$ .

This follows because when injurers are risk neutral, the right-hand side of (9) reduces to  $u - r + v - q(p(x_s)h + x_s)$ , which is maximized when  $p(x_s)h + x_s$  is minimized over  $x_s$ .

We now have the following result.

*Proposition 1.* Assume that injurers are risk averse and that victims are risk neutral. Then the Pareto optimal standard of care  $x_s^*$  is such that  $0 < x_s^* < x^*$ , where  $x^*$  is the conventionally optimal level of care. The Pareto optimal standard  $x_s^*$  is determined by (14) below.

*Notes.* (a) The explanation for why  $x_s^*$  is positive is as follows. If  $x_s$  were zero, injurers would bear no risk — if the activity was discovered to be harmful, the state would impose no standard. But since a risk-averse party is effectively risk neutral when he begins to bear risk, we infer that it would be desirable for  $x_s$  to be raised slightly from zero if that would be desirable for a risk-neutral injurer. For a risk-neutral injurer, it would be desirable for  $x_s$  to be raised marginally from zero, since we know from Remark 1 that it would be desirable for  $x_s$  to minimize  $p(x_s)h + x_s$  that for such an injurer. Hence, it must be desirable for  $x_s$  to be raised slightly from zero for a risk-averse injurer. In other words, doing so will lower the required payment  $t$  by enough to make the cost of meeting  $x_s$  worthwhile for him.

(b) The explanation for why  $x_s^*$  is below  $x^*$  flows from the observation that if  $x_s$  is marginally reduced from  $x^*$ , then because  $x^*$  minimizes  $x + p(x)h$ , the increase in  $t$  will essentially equal the expected reduction in  $x_s$ .<sup>26</sup> Thus, if the injurer were risk neutral, the

---

<sup>24</sup> That is, maximization of (4) subject to (5) is maximization of (6) subject to (7) in the present context.

<sup>25</sup> Since  $EU''(x_s) < 0$  will be shown in (11) below,  $x_s^*$  is unique.

<sup>26</sup> Since  $t = r - v + qp(x^*)h$ , the marginal increase in  $t$  from a marginal reduction in  $x_s$  is  $-qp'(x^*)h$ . Since the expected cost of meeting the standard is  $qx^*$ , the marginal reduction in expected costs is  $q$ . And since, by (2),  $x^*$  satisfies  $p'(x^*)h = -1$ , we know that  $-qp'(x^*)h = q$ .

marginal effect of the reduction in  $x_s$  on his well-being would be zero. But since the injurer is risk averse and is bearing positive risk when  $x_s$  is  $x^*$ , he benefits from a reduction in risk-bearing from the marginal reduction in  $x_s$ , implying that his expected utility must rise.

*Proof.* First, observe that

$$(10) \quad EU'(x_s) = -qp'(x_s)h(1-q)U'(w - qp(x_s)h) - (qp'(x_s)h + 1)qU'(w - qp(x_s)h - x_s),$$

where  $w$  denotes  $u - r + v$  for notational simplicity. Hence

$$(11) \quad EU''(x_s) = -qp''(x_s)h(1-q)U'(w - qp(x_s)h) + (qp'(x_s)h)^2(1-q)U''(w - qp(x_s)h) \\ - qp''(x_s)hqU'(w - qp(x_s)h - x_s) + (qp'(x_s)h + 1)^2qU''(w - qp(x_s)h - x_s) < 0,$$

since each term is negative.

To show that  $x_s^* > 0$ , note from (10) that

$$(12) \quad EU'(0) = -qp'(0)h(1-q)U'(w - qp(0)h) - (qp'(0)h + 1)qU'(w - qp(0)h) \\ = -q(p'(0)h + 1)U'(w - qp(0)h) > 0$$

since  $p'(0)h + 1 < 0$  by (3).

To show that  $x_s^* < x^*$ , it suffices to show that  $EU'(x^*) < 0$ , since  $EU''(x_s) < 0$ . Now

$$(13) \quad EU'(x^*) = -qp'(x^*)h(1-q)U'(w - qp(x^*)h) - (qp'(x^*)h + 1)qU'(w - qp(x^*)h - x^*) \\ < -qp'(x^*)h(1-q)U'(w - qp(x^*)h - x^*) - (qp'(x^*)h + 1)qU'(w - qp(x^*)h - x^*) \\ = [-qp'(x^*)h(1-q) - (qp'(x^*)h + 1)q]U'(w - qp(x^*)h - x^*) \\ = -q(p'(x^*)h + 1)U'(w - qp(x^*)h - x^*) = 0,$$

since  $p'(x^*)h + 1 = 0$  by (2).

The first-order condition determining  $x_s^*$  is

$$(14) \quad -qp'(x_s)h[(1-q)U'(w - qp(x_s)h) + qU'(w - qp(x_s)h - x)] = qU'(w - qp(x_s)h - x_s).$$

On the left is the marginal benefit to an injurer of raising the standard; it is the amount  $-qp'(x_s)h$  by which the payment to victims falls, weighted by the expected marginal utility of money. On the right is the marginal cost of raising the standard, which involves a high marginal utility of money because the cost is incurred when the standard is imposed.<sup>27</sup> Q. E. D.

Let us next consider the case where victims are risk averse. Then the expected utility of a victim is

$$(15) \quad EV(x_s) = (1-q)V(v+t) + qV(v+t-p(x_s)h).$$

The risk borne by the victim is the expected harm given the standard,  $p(x_s)h$ , because it is assumed that a victim pays the actuarially fair premium for insurance coverage against suffering a loss  $h$ .<sup>28</sup> The Pareto optimal standard is therefore determined by maximizing (6) subject to

$$(16) \quad (1-q)V(v+t) + qV(v+t-p(x_s)h) = r,$$

where  $r$  is a reference level of expected utility of victims.

<sup>27</sup> I will not comment on first-order conditions in later propositions because their interpretations are similar.

<sup>28</sup> If the premium for insurance coverage is actuarially fair, it is a standard result that risk-averse parties would maximize their expected utility by purchasing full coverage; see, for example, Shavell (1987), chapter 8. In any event, even if victims cannot insure against  $h$ , the qualitative conclusion I draw in the next proposition would not change.

I wish to show that an analogue of Proposition 1 holds. A way to view Proposition 1 is that it shows that when victims are risk neutral, the effect of injurer risk aversion is to lower the Pareto optimal standard from what it would otherwise be (for when injurers are risk neutral, the Pareto optimal standard is  $x^*$  by Remark 1). Hence, a natural analogue of Proposition 1 is that when victims are risk averse, the effect of injurer risk aversion is also to lower the Pareto optimal standard from what it would otherwise be. This is demonstrated in the next result.

*Proposition 2.* Assume that injurers are risk averse and that victims are also risk averse. Then the Pareto optimal standard of care  $x_s^*$  is such that  $0 < x_s^* < x^{**}$ , where  $x^{**}$  is the Pareto optimal standard of care when injurers are risk neutral and victims are risk averse. Here  $x^* < x^{**}$ , and  $x_s^*$  is determined by (27) below.

*Notes.* (a) The explanation for why  $x_s^*$  is positive is similar to that given in Proposition 1. The only difference is that here, because victims are risk averse, the reduction in the payment  $t$  that injurers make if  $x_s$  is raised slightly from zero is greater than when victims are risk neutral, for risk averse victims benefit from a reduction in risk bearing. This reinforces the benefit to injurers from raising  $x_s$  from zero.

(b) The explanation for why  $x_s^* < x^{**}$  is analogous to that given in Proposition 1 for why  $x_s^* < x^*$ . Namely, if  $x_s$  is marginally reduced from  $x^{**}$ , then since the injurer is risk averse and is bearing positive risk, he benefits from a reduction in risk-bearing, whereas this benefit is not reflected in the determination of  $x^{**}$ .

(c) The explanation for why  $x^* < x^{**}$  is that when victims are risk averse, they benefit from higher  $x_s$  not only because care lowers expected harm but also because care lowers risk-bearing. Thus, the reduction in  $t$  is greater when  $x_s$  is raised. Accordingly, risk-neutral injurers find a higher  $x_s$  desirable than when victims are risk neutral.

*Proof.* Equation (16) determines  $t$  as a function of  $x_s$ , which we write as  $t(x_s)$ . Implicit differentiation of (16) with respect to  $x_s$  gives

$$(17) \quad t'(x_s) = qp'(x_s)hz(x_s) < 0,$$

where

$$(18) \quad z(x_s) = V'(v + t(x_s) - p(x_s)h) / [(1 - q)V'(v + t(x_s)) + qV'(v + t(x_s) - p(x_s)h)].$$

Note that

$$(19) \quad z(x_s) > 1.$$

Differentiation of (17) gives

$$(20) \quad t''(x_s) = qp''(x_s)hz(x_s) + qp'(x_s)hz'(x_s).$$

Since  $z'(x_s) < 0$ ,<sup>29</sup> (20) implies that

---

<sup>29</sup> It suffices to demonstrate that the numerator of the derivative of  $z(x_s)$  is negative since the denominator must be positive. This numerator is readily verified to be  $(1 - q)(t'(x_s) - p'(x_s)h)V''(v + t - p(x_s)h)V'(v + t) - (1 - q)V'(v + t - p(x_s)h)t'(x_s)V''(v + t)$ . Since the second term is negative, we need to show that the first term is negative. The latter will be true if  $t'(x_s) - p'(x_s)h > 0$ . But  $t'(x_s) - p'(x_s)h = qp'(x_s)hz(x_s) - p'(x_s)h = p'(x_s)h(qz(x_s) - 1)$ . Because it is clear that  $qz(x_s) < 1$ , it must be that  $t'(x_s) - p'(x_s)h > 0$  as required.

$$(21) \quad t''(x_s) > 0.$$

Let us first show that  $x_s^* > 0$ . Now  $x_s^*$  is determined by maximizing

$$(22) \quad EU(x_s) = (1 - q)U(u - t(x_s)) + qU(u - t(x_s) - x_s)$$

over  $x_s$ . We have

$$(23) \quad EU'(x_s) = -(1 - q)t'(x_s)U'(u - t(x_s)) - (t'(x_s) + 1)qU'(u - t(x_s) - x_s).$$

Hence,

$$(24) \quad \begin{aligned} EU'(0) &= -(1 - q)t'(0)U'(u - t(0)) - (t'(0) + 1)qU'(u - t(0)) \\ &= -(t'(0) + q)U'(u - t(0)). \end{aligned}$$

But

$$(25) \quad t'(0) = qp'(0)hz(0) < qp'(0)h < -q,$$

where the first inequality holds because  $z(0) > 1$  and the second inequality holds because of (3).

Hence,  $EU'(0) > 0$ , so that  $x_s^* > 0$  as claimed.

Let us next prove that  $x_s^* < x^{**}$ . The latter is found by maximizing

$$(26) \quad \begin{aligned} EU(x_s) &= (1 - q)U(u - t(x_s)) + qU(u - t(x_s) - x_s) \\ &= (1 - q)(u - t(x_s)) + q(u - t(x_s) - x_s) = u - (t(x_s) + qx_s) \end{aligned}$$

over  $x_s$  since injurers are risk neutral. Therefore  $x^{**}$  is determined by

$$(27) \quad t'(x_s) = -q.$$

Given (27) and the fact that  $t''(x_s) > 0$ , we know that if  $t'(x_s^*) < -q$ , it must be that  $x_s^* < x^{**}$ .

Hence, we want to show  $t'(x_s^*) < -q$ . From (23), we know that  $x_s^*$  is determined by

$$(28) \quad -(1 - q)t'(x_s)U'(u - t(x_s)) - (t'(x_s) + 1)qU'(u - t(x_s) - x_s) = 0,$$

which implies that at  $x_s^*$ ,

$$(29) \quad \begin{aligned} t'(x_s) &= -qU'(u - t(x_s) - x_s)/[(1 - q)U'(u - t(x_s)) + qU'(u - t(x_s) - x_s)] \\ &< -qU'(u - t(x_s) - x_s)/[(1 - q)U'(u - t(x_s) - x_s) + qU'(u - t(x_s) - x_s)] = -q, \end{aligned}$$

establishing the claim.

Last, let us show that  $x^* < x^{**}$ . From (27) and (17), we have

$$(30) \quad qp'(x^{**})hz(x^{**}) = -q,$$

which implies that

$$(31) \quad p'(x^{**})h = -1/z(x^{**}) > -1$$

since, by (19),  $z(x^{**}) > 1$ . Because  $p'(x^*)h = -1$  and  $p''(x) > 0$ , (31) implies

that  $x^{**} > x^*$ . Q. E. D.

### 3.2 Payment for Harm

Now assume that if the state learns that the activity is harmful, the legal rule that the state adopts requires injurers to make payments to victims based on harm. Two rules of this type that will be equivalent under our assumptions will be considered: strict liability and corrective taxes.

Under strict liability, injurers must pay “damages” to victims if harm occurs.<sup>30</sup> Let

$d$  = damages payment that an injurer must make if he causes harm.

---

<sup>30</sup> A fine paid to the state would be equivalent to damages paid to victims given the assumptions in the model.

In our legal system,  $d$  is generally intended to equal  $h$ , but here we will determine the Pareto optimal  $d$ . Under strict liability, it will be assumed that, if the injurer is risk averse and strict liability is imposed, the injurer can purchase liability insurance at an actuarially fair rate. Hence, if strict liability is imposed, the injurer will purchase full coverage against  $d$  and pay a premium of  $p(x)d$ .<sup>31</sup> (Note that the assumption that injurers can buy liability insurance if the activity is found harmful is consistent with the assumption that injurers cannot buy coverage against legal change, that is, the decision of the state to use of the rule of strict liability. In any event, the assumption is not essential.<sup>32</sup>) Hence, the injurer's utility will be

$$(32) \quad EU = (1 - q)U(u - t) + qU(u - t - x - p(x)d).$$

Under corrective taxes, injurers are required to make a payment of  $p(x)d$  if the state learns that the activity is harmful. Hence, if  $d$  equals  $h$ , the payment equals the expected harm from the activity. The expected utility of an injurer under the corrective tax is also given by (32).

Note that the risk borne by the injurer is the cost of care  $x$  plus the payment  $p(x)d$  (whereas under regulation the risk was only the cost of care).

The level of care that an injurer chooses maximizes (32),<sup>33</sup> so that it minimizes

$$(33) \quad x + p(x)d.$$

Since (33) is strictly convex in  $x$ , there exists a unique  $x$  minimizing it, which will be denoted  $x(d)$ . Further, since the derivative of (33) is  $1 + p'(x)d$ , we know that if  $1 + p'(0)d \geq 0$ , then  $x(d) = 0$ . Consequently, if  $d \leq -1/p'(0)$ , then  $x(d) = 0$ . Otherwise,  $x(d)$  is positive and is determined by

$$(34) \quad 1 + p'(x)d = 0.$$

From implicit differentiation of (34), we obtain that  $p''(x)x'(d)d + p'(x) = 0$ , implying that

$$(35) \quad x'(d) = -p'(x)/(p''(x)d) > 0.$$

If victims are risk neutral, the expected wealth of a victim will be  $v + t - qp(x)(h - d)$ ,<sup>34</sup> since victims bear  $h - d$  of their losses.<sup>35</sup> Thus, we must have

$$(36) \quad v + t - qp(x)(h - d) = r,$$

where  $r$  is again a reference level of wealth. Hence,

$$(37) \quad t = r - v + qp(x)(h - d).$$

---

<sup>31</sup> This presumes that insurers can observe  $x$  and base the premium on it; see, for example, Shavell (1987), at chapter 8.

<sup>32</sup> If instead it was assumed that injurers could not purchase liability insurance, the conclusion to be shown that  $d < h$  is Pareto optimal would only be reinforced, as injurer risk-bearing due to legal uncertainty would be greater.

<sup>33</sup> It is supposed that injurers take  $t$  as given, the motivation for which is that there are many injurers so that actions of any single injurer would have only a negligible effect on  $t$ .

<sup>34</sup> Here and below, I will sometimes write  $x$  instead of  $x(d)$  for notational convenience.

<sup>35</sup> Under the rule of strict liability, victims are directly compensated by injurers. Under corrective taxes, it is the government that collects payments and I will interpret the government as the victim.

The Pareto optimal level of damages  $d$  for risk-averse injurers and risk neutral victims thus maximizes (32) subject to (37). Hence,  $d^*$  is determined by maximization of

$$(38) \quad EU(d) = (1 - q)U(u - r + v - qp(x(d))(h - d)) \\ + qU(u - r + v - qp(x(d))(h - d) - x(d) - p(x(d))d).$$

Before we determine this  $d^*$ , let us note the following.

*Remark 2.* Assume that both injurers and victims are risk neutral. Then the Pareto optimal level of damages  $d^*$  is the harm  $h$ .

Remark 2 confirms that when parties are risk neutral, the conventionally optimal level of liability equal to the harm is Pareto optimal and that the conventionally best corrective tax equal to the expected harm is Pareto optimal. To demonstrate these conclusions, observe that when injurers are risk neutral, (38) reduces to

$$(39) \quad u - r + v - q[p(x(d))h + x(d)].$$

To maximize (39),  $[p(x(d))h + x(d)]$  must be minimized over  $d$ . That is achieved if  $d$  is  $h$ , for by definition,  $x(h)$  minimizes  $p(x)h + x$  over  $x$ .

Now let me describe the result when injurers are risk averse. We have

*Proposition 3.* Assume that injurers are risk averse and that victims are risk neutral. Then the Pareto optimal level of damages  $d^*$  is such that  $0 \leq d^* < h$ , where  $h$  is harm; the Pareto optimal corrective tax is  $p(x(d^*))d^*$  and thus satisfies  $0 \leq p(x(d^*))d^* < p(x(d^*))h$ . Further,  $d^* = 0$  if  $h$  is in  $[0, \hat{h}]$ , where  $\hat{h} > -1/p'(0)$ , and if  $d^* > 0$ , it is determined by (47) below.

*Notes.* (a) The explanation for why  $d^*$  will be zero for  $h$  sufficiently low is as follows. Raising  $d$  from zero has no effect on  $x$  until  $d$  exceeds  $-1/p'(0)$ , as was noted above. In other words, a positive level of risk must be imposed on injurers to induce them to begin to raise  $x$  and lower  $p(x)$ . This imposition of risk-bearing may not be worthwhile because it may exceed the benefit to injurers from lowering the payment  $t$  that they make to victims.<sup>36</sup> That  $\hat{h} > -1/p'(0)$  implies that there are  $h$  for which  $x^*(h)$  is positive — the conventionally optimal level of care is positive—yet for which  $d^*$  is zero due to factor of risk-bearing.

(b) The explanation for why  $d^* < h$  is that if  $d$  is slightly lowered from  $h$ , then the increase in  $t$  will approximate the expected reduction in care and payments,<sup>37</sup> that is, if the

---

<sup>36</sup> Note the contrast with Proposition 1, where the optimal standard of care  $x_s^*$  must be positive. When the level of care can be directly controlled as a standard by the state, no risk is imposed on injurers when the standard begins to be raised from zero. But here, as explained, when the level of care is only indirectly controlled by the state through imposition of payments for harm, positive risk must be imposed before the standard begins to be raised from zero. This explains the difference in conclusions.

<sup>37</sup> Since  $t = r - v + qp(x(d))(h - d)$ , the increase in  $t$  from a marginal reduction in  $d$  is

injurer were risk neutral, the marginal effect on his well-being would be zero. But since the injurer is risk averse and is bearing positive risk, he benefits from a reduction in risk-bearing from the marginal reduction in  $d$ , implying that his expected utility must rise.

*Proof.* I first show that  $EU(d)$  is strictly decreasing in  $d$  for  $d$  in  $[0, -1/p'(0)]$ . From (38), we have

$$(40) \quad EU'(d) = (1 - q)[-qp'(x)x'(d)(h - d) + qp(x)]U'(u - r + v - qp(x(d))(h - d)) \\ + q[-qp'(x)x'(d)(h - d) + qp(x) - x'(d) - p'(x)x'(d)d - p(x)] \\ \times U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d.$$

Now for  $d$  in  $[0, -1/p'(0)]$ , recall that  $x(d) = 0$ . Hence,  $x'(d) = 0$  in the interval. Thus, in that interval (40) reduces to

$$(41) \quad EU'(d) = (1 - q)[qp(0)]U'(u - r + v - qp(0)(h - d)) \\ + q[qp(0) - p(0)]U'(u - r + v - qp(0)(h - d)) - p(0)d.$$

Hence, for  $d$  in  $(0, -1/p'(0)]$

$$(42) \quad EU'(d) < (1 - q)[qp(0)]U'(u - r + v - qp(0)(h - d)) - p(0)d \\ + q[qp(0) - p(0)]U'(u - r + v - qp(0)(h - d)) - p(0)d = 0.$$

Let us next show that for any  $h$  in  $[0, -1/p'(0)]$ ,  $EU(d)$  is strictly decreasing in  $d$  for  $d > -1/p'(0)$  and also that  $d^*(h) = 0$ . For  $d > -1/p'(0)$ , we know that  $x(d) > 0$ , and from (40) we see that

$$(43) \quad EU'(d) < \{(1 - q)[-qp'(x)x'(d)(h - d) + qp(x)] \\ + q[-qp'(x)x'(d)(h - d) + qp(x) - x'(d) - p'(x)x'(d)d - p(x)]\} \\ \times U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d \\ = \{[-qp'(x)x'(d)(h - d)] - q[x'(d) + p'(x)x'(d)d]\} \\ \times U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d.$$

But  $-qp'(x)x'(d)(h - d) < 0$  since  $d > h$  when  $d > -1/p'(0)$  and  $h$  is in  $[0, -1/p'(0)]$ ; and  $[x'(d) + p'(x)x'(d)d] = x'(d)(1 + p'(x)d) = 0$  since  $x(d)$  satisfies  $1 + p'(x)d = 0$ . Thus, the last line of (43) is negative, showing that  $EU'(d) < 0$  for  $d > -1/p'(0)$ . Consequently,  $EU(d)$  is decreasing for all  $d$  given that  $h$  is in  $[0, -1/p'(0)]$ . It thus follows that  $d^*(h) = 0$  for such  $h$ .

We now show that for  $h > -1/p'(0)$  and sufficiently close to  $-1/p'(0)$ ,  $d^*(h) = 0$ , so that the asserted  $\hat{h}$  exists. We do this using a number of observations. (i) For any  $h$ ,  $d^*(h)$  is either 0 or exceeds  $-1/p'(0)$ : This is true since  $EU(d)$  is strictly decreasing in  $d$  in  $[0, -1/p'(0)]$ , as demonstrated in the first paragraph in this proof. (ii) Let  $m(h) = \text{maximum of } EU(d, h) \text{ over } d \geq -1/p'(0)$ , where  $EU(d, h)$  is  $EU(d)$  given  $h$ . Then  $m(h)$  is clearly decreasing in  $h$ . (iii)  $m(-1/p'(0)) > m(h)$  for  $h > -1/p'(0)$ : This follows from (ii). (iv)  $EU(0, h) = U(u - r + v - qp(0)h)$  is continuous in  $h$ . (v)  $EU(0, -1/p'(0)) > EU(-1/p'(0), -1/p'(0)) = m(-1/p'(0))$ : The inequality here follows because we showed that  $EU(d)$  is decreasing for  $d$  in  $[0, -1/p'(0)]$ , and the equality

---

$-qp'(x)x'(d)(h - d) + qp(x)$ , so that at  $d = h$ , the marginal increase in  $t$  is  $qp(x(h))$ . Expected care is  $qx(d)$ , so the marginal reduction in this quantity at  $d = h$  is  $-qx'(h)$ . Expected payments are  $qp(x(d))d$ , so the marginal reduction in these at  $d = h$  is  $-qp'(x(h))x'(h)h - qp(x(h))$ . Hence, the change in expected care and payments is  $-qx'(h) - qp'(x(h))x'(h)h - qp(x(h)) = -qp(x(h))$  since  $qx'(h) + qp'(x(h))x'(h)h = qx'(h)(1 + p'(x(h))h) = 0$ . Thus the marginal increase in  $t$  does indeed offset the marginal decrease in expected expenses.

follows because we showed that  $EU(d)$  is decreasing for larger  $d$  when  $h$  does not exceed  $-1/p'(0)$ . (vi)  $EU(0, h) > m(-1/p'(0))$  for all  $h$  above  $-1/p'(0)$  and sufficiently close to it: This follows from (iv) and (v). (vii)  $EU(0, h) > m(h)$  for all  $h$  above  $-1/p'(0)$  and sufficiently close to it: This follows from (iii) and (vi). (viii)  $d^*(h) = 0$  for all  $h$  above  $-1/p'(0)$  and sufficiently close to it: This follows from (vii) and (i).

To show that  $d^* < h$ , rewrite (40), making use of the fact that  $x'(d)(1 + p'(x)d) = 0$  since  $x(d)$  satisfies  $1 + p'(x)d = 0$ , to obtain

$$(44) \quad EU'(d) = (1 - q)[-qp'(x)x'(d)(h - d) + qp(x)]U'(u - r + v - qp(x(d))(h - d)) \\ + q[-qp'(x)x'(d)(h - d) + qp(x) - p(x)]U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d.$$

We want to show that (44) is negative for  $d \geq h$ . Since  $-qp'(x)x'(d)(h - d) \leq 0$  for  $d \geq h$ , it suffices to show that

$$(45) \quad (1 - q)qp(x)U'(u - r + v - qp(x(d))(h - d)) \\ + q[qp(x) - p(x)]U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d < 0$$

for  $d \geq h$ . But the left-hand side equals

$$(46) \quad qp(x)(1 - q) \\ \times [U'(u - r + v - qp(x(d))(h - d)) - U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d] < 0.$$

The first-order condition determining  $d^*$  is, from (44),

$$(47) \quad [-qp'(x)x'(d)(h - d) + qp(x)] \\ \times [(1 - q)U'(u - r + v - qp(x(d))(h - d)) + qU'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d] \\ = qp(x)U'(u - r + v - qp(x)(h - d)) - x(d) - p(x)d.$$

Q. E. D.

Let us next consider the case in which victims are risk averse. Hence, the expected utility of a victim is

$$(48) \quad EV(d) = (1 - q)V(v + t) + qV(v + t - p(x(d))(h - d))$$

since victims pay the fair insurance premium  $p(x(d))(h - d)$  for coverage against the risk  $h - d$  that they bear. The Pareto optimal standard is therefore determined by maximizing (32) subject to

$$(49) \quad (1 - q)V(v + t) + qV(v + t - p(x(d))(h - d)) = r,$$

where  $r$  is a reference level of expected utility of victims. We then have

*Proposition 4.* Assume that injurers are risk averse and that victims are also risk averse. Then the Pareto optimal level of damages  $d^*$  is such that  $0 < d^* < h$ , where  $h$  is harm; the Pareto optimal corrective tax is  $p(x(d^*))d^*$  and satisfies  $0 < p(x(d^*))d^* < p(x(d^*))h$ . Further,  $d^*$  is determined by (61) below.

*Notes.* (a) The explanation for why  $d^*$  is positive follows from the point that  $x(d) = 0$  for  $d$  in  $[0, -1/p'(0)]$ . In other words, raising  $d$  in this interval shifts risk from injurers to victims but does not change the probability of harm and thus the risk to be shared between them. Since

some degree of risk-sharing of a given risk between two risk-averse parties is desirable,  $d^*$  must be positive.<sup>38</sup>

(b) The explanation for why  $d^* < h$  follows from the explanation of this result in the previous proposition, where victims were risk-neutral. Here, although victims are risk averse, they behave essentially as if they were risk neutral when  $d = h$ , for then they bear no risk.

*Proof.* Equation (49) determines  $t(d)$ , and differentiating (49) with respect to  $d$  gives

$$(50) \quad (1 - q)t'(d)V'(v + t(d)) + q[t'(d) - p'(x)x'(d)(h - d) + p(x(d))]V'(v + t(d) - p(x(d))(h - d)) = 0.$$

Hence,

$$(51) \quad t'(d) = q[p'(x)x'(d)(h - d) - p(x(d))] \times V'(v + t(d) - p(x(d))(h - d)) / [(1 - q)V'(v + t(d)) + qV'(v + t(d) - p(x(d))(h - d))].$$

The optimal  $d$  is determined by maximizing

$$(52) \quad EU(d) = (1 - q)U(u - t(d)) + qU(u - t(d) - x(d) - p(x(d))d),$$

the derivative of which is

$$(53) \quad EU'(d) = -(1 - q)t'(d)U'(u - t(d)) - q[t'(d) + x'(d) + p'(x)x'(d)d + p(x(d))]U'(u - t(d) - x(d) - p(x(d))d).$$

Since  $x(d) = 0$  for  $d$  in  $[0, -1/p'(0)]$ ,  $x'(d) = 0$  in this interval. Hence,

$$(54) \quad t'(0) = -qp(0)V'(v + t(0) - p(0)h) / [(1 - q)V'(v + t(0)) + qV'(v + t(0) - p(0)h)]$$

and

$$(55) \quad EU'(0) = -(1 - q)t'(0)U'(u - t(0)) - q[t'(0) + p(0)]U'(u - t(0)) = -[t'(0) + qp(0)]U'(u - t(0)).$$

But from (54),

$$(56) \quad t'(0) + qp(0) = qp(0) \{ 1 - V'(v + t(0) - p(0)h) / [(1 - q)V'(v + t(0)) + qV'(v + t(0) - p(0)h)] \} < 1.$$

Thus,  $EU'(0) > 0$ , showing that  $d^* > 0$ .

To demonstrate that  $d^* < h$ , it is sufficient to show that  $EU'(d) < 0$  for  $d \geq h$ . We know that  $x'(d) + p'(x)x'(d)d = 0$ , since this must be true for  $d$  in  $[0, -1/p'(0)]$  as was noted in the previous paragraph and since for greater  $d$ ,  $x(d) > 0$ , so that  $1 + p'(x(d))d = 0$ . Consequently, (53) reduces to

$$(57) \quad EU'(d) = -(1 - q)t'(d)U'(u - t(d)) - q[t'(d) + p(x(d))]U'(u - t(d) - x(d) - p(x(d))d) = -t'(d)[(1 - q)U'(u - t(d)) + qU'(u - t(d) - x(d) - p(x(d))d)] - qp(x(d))U'(u - t(d) - x(d) - p(x(d))d).$$

To show that (57) is negative for  $d \geq h$ , let us first verify that  $t'(d) > -qp(x(d))$  for  $d \geq h$ .

Using (51), the latter inequality is equivalent to

$$(58) \quad q[p'(x)x'(d)(h - d) - p(x(d))] \times V'(v + t(d) - p(x(d))(h - d)) / [(1 - q)V'(v + t(d)) + qV'(v + t(d) - p(x(d))(h - d))] \geq -qp(x(d)).$$

Since  $p'(x)x'(d)(h - d) \geq 0$  when  $d \geq h$ , (58) will be true if

---

<sup>38</sup> In Proposition 3, this logic did not apply since victims were risk neutral.

$$(59) -qp(x(d)) \times \{V'(v + t(d) - p(x(d))(h - d))/[(1 - q)V'(v + t(d)) + qV'(v + t(d) - p(x(d))(h - d))]\} \geq -qp(x(d)).$$

This must hold, for the term in braces is at least 1.

Since  $t'(d) \geq -qp(x(d))$  or  $-t'(d) \leq qp(x(d))$  for  $d \geq h$ , we have from (57) for such  $d$  that

$$(60) EU'(d) \leq qp(x(d))[(1 - q)U'(u - t(d)) + qU'(u - t(d) - x(d) - p(x(d))d)] - qp(x(d))U'(u - t(d) - x(d) - p(x(d))d) = qp(x(d))\{(1 - q)U'(u - t(d)) + qU'(u - t(d) - x(d) - p(x(d))d)\} - U'(u - t(d) - x(d) - p(x(d))d) < 0.$$

Hence,  $d^* < h$ .

The first-order condition determining  $d^*$  is

$$(61) -t'(d)[(1 - q)U'(u - t(d)) + qU'(u - t(d) - x(d) - p(x(d))d)] = qp(x(d))U'(u - t(d) - x(d) - p(x(d))d).$$

Q.E.D.

#### 4. Concluding Comments

(a) *The generality of the model.* The chief qualitative conclusion of the analysis — that legal change should be attenuated due to the bearing of risk by risk-averse parties subject to the law — does not depend on a number of simplifying features of the model. One assumption was that there was a single type of action parties could take, the exercise of care after a legal change. If other types of action were considered, notably, decisions about levels of activity after a legal change, or decisions about levels of care or of activity before a legal change, nothing essential would be altered. There would then be other dimensions of inefficiency of behavior engendered by the attenuation of legal change, but attenuation would still be generally desirable to relieve risk-bearing. A second simplifying assumption was that the source of the change in the law was the state's learning that an activity was harmful rather than harmless. If the source of the change was, instead, information about the level of harmfulness of an activity already known to be harmful, or about the technology or cost of harm reduction, that would obviously not alter the main conclusion. A third simplifying assumption was that there was no insurance available against legal change. If partial insurance were available despite the problem of the correlation of losses, risk-averse parties would still bear some risk, so that attenuation of the legal change might again be desirable.

The importance of the conclusion that attenuation of legal change is beneficial is a function of the degree of risk aversion of parties subject to the law and of the magnitude of the risks that they bear. Thus for individuals facing substantial risks, such as homeowners with large mortgages confronting the risk of elimination of their interest deduction, significant attenuation is presumably desirable, whereas for publicly-held firms facing the risk of a modest change in a workplace safety rule, little if any adjustment of the rule from efficiency would be appropriate.

(b) *Attenuation of legal change in reality.* Legal change often appears to be attenuated in practice, reflecting the hardships that it could cause for the parties to whom it is addressed. If, for example, a proposed regulation aimed at reduction of pollution would impose a large expense on small businesses, they might be able to exert their influence to lessen the rigor of the

regulation. Moreover, legal change is frequently delayed or implemented in phases. A 10% increase in a tax rate might be achieved through a 5% increase in one year and another 5% increase in a second year. Such implementation by its nature involves attenuation, for parties subject to the change are relieved from its full effects for a period of time. Also of relevance is grandfathering, under which parties who are already engaged in an activity governed by a legal rule are excused from having to satisfy a change to it. Grandfathering is obviously a form of attenuation, as it arrests change for those engaged in an activity. In sum, legal change seems to display a broadly attenuated character in reality, which is consistent with the analysis of this article.<sup>39</sup>

---

<sup>39</sup> Although I have emphasized the reduction of risk-bearing as a reason for attenuation of legal change here, another rationale is avoidance of the waste of past expenditures made to comply with the law, the theme developed in Shavell (2008). In addition, attenuation of the law may be ascribed simply to the self-interest and political power of parties who would be adversely affected by a legal change — whether or not these parties are risk averse.

## References

- Abraham, Kenneth S. 1977. "Medical Malpractice Reform: A Preliminary Analysis," 36 *Maryland Law Review* 489-532.
- Abraham, Kenneth S. 2010. *Insurance Law and Regulation*. Fifth edition. New York: Foundation Press.
- Appleman on Insurance, 2d*. 1996. Second edition. St. Paul: West.
- Baker, Tom. 2004. "Insuring Liability Risks," 29 *The Geneva Papers on Risk and Insurance* 128-149.
- Baker, Tom. 2008. *Insurance Law and Policy*. Second edition. New York: Aspen.
- Blume, Lawrence, and Daniel L. Rubinfeld. 1984. "Compensation for Takings: An Economic Analysis," 72 *California Law Review* 569-628.
- Church, George J. 2005. "Sorry, Your Policy Is Canceled," *Time Magazine*. Tuesday, June 21, 2005. Available at <<http://www.time.com/time/printout/0,8816,1075044,00.html>>
- Couch on Insurance, 3d*. 2009. Revised edition. St. Paul: West.
- Froot, Kenneth A. 2001. "The Market for Catastrophe Risk: A Clinical Examination," 60 *Journal of Financial Economics* 529-571.
- Graetz, Michael. 1977. "Legal Transitions: The Case of Retroactivity in Income Tax Revision," 126 *University of Pennsylvania Law Review* 47-87.
- Harrington, Scott E., and Gregory R. Niehaus. 2004. *Risk Management and Insurance*. Second edition. New York: McGraw-Hill/Irwin.
- Kaplow, Louis. 1986. "An Economic Analysis of Legal Transitions," 99 *Harvard Law Review* 509-617.
- Kaplow, Louis. 1992. "Government Relief for Risk Associated with Government Action," 94 *Scandinavian Journal of Economics* 525-541.
- Landes, William M., and Richard A. Posner. 1987. *The Economic Structure of Tort Law*. Cambridge, MA: Harvard University Press.

- Levmore, Saul. 1998. Retroactive Taxation. 3:340–44 in *The New Palgrave Dictionary of Economics and the Law*, edited by Peter Newman. New York: Stockton Press.
- Levmore, Saul. 1999. “Changes, Anticipations, and Reparations,” 99 *Columbia Law Review* 1657–1700.
- Logue, Kyle D. 1996. “Tax Transitions, Opportunistic Retroactivity, and the Benefits of Government Precommitment,” 94 *Michigan Law Review* 1129–1196.
- Logue, Kyle D. 2003. “Legal Transitions, Rational Expectations, and Legal Progress,” 13 *Journal of Law and Contemporary Legal Issues* 211–260.
- Masur, Jonathan S., and Jonathan Remy Nash. 2010. “The Institutional Dynamics of Transition Relief,” 85 *New York University Law Review* 391–456.
- Nash, Jonathan Remy, and Richard L. Revesz. 2007. “Grandfathering and Environmental Regulation: The Law and Economics of New Source Review,” 101 *Northwestern University Law Review* 1677–1733.
- Priest, George L. 1987. “The Current Insurance Crisis and Modern Tort Law,” 96 *Yale Law Journal* 1521–1590.
- Quinn, John, and Michael J. Trebilcock. 1982. “Compensation, Transition Costs, and Regulatory Change,” 32 *University of Toronto Law Journal* 117–175.
- Shavell, Steven. 1987. *Economic Analysis of Accident Law*. Cambridge, MA: Harvard University Press.
- Shavell, Steven. 2008. “On Optimal Legal Change, Past Behavior, and Grandfathering,” 37 *Journal of Legal Studies* 37–85.
- Shaviro, Daniel. 2000. *When Rules Change. An Economic and Political Analysis of Transition Relief and Retroactivity*. Chicago: University of Chicago Press.
- Vaughan, Emmett J., and Therese Vaughan. 2008. *Fundamentals of Risk and Insurance*. Tenth Edition. Hoboken, NJ: John Wiley.
- Viscusi, W. Kip. 1991. “The Dimensions of the Product Liability Crisis,” 20 *Journal of Legal Studies* 147–177.

Wallace, Charles. 2011. "Is it Time to Kill the Mortgage Interest Tax Deduction?" *Daily Finance*. April 20 2011, available at <http://www.dailyfinance.com/2011/04/20/eliminate-mortgage-interest-tax-deduction/>.