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

HARVARD

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

RISK AVERSION AND THE DESIRABILITY OF ATTENUATED LEGAL CHANGE

Steven Shavell

Discussion Paper No. 756

01/2014

Forthcoming in the *American Law and Economics Review*

Harvard Law School
Cambridge, MA 02138

Risk Aversion and the Desirability of Attenuated Legal Change

Steven Shavell

forthcoming in the *American Law and Economics Review*

Risk Aversion and the Desirability of Attenuated Legal Change

Steven Shavell*

This article develops two points. First, insurance against the risk of legal change is largely unavailable, primarily because of the correlated nature of the losses that legal change generates. Second, given the absence of insurance against legal change, it is generally desirable for legal change to be attenuated. Specifically, in a model of uncertainty about two different types of legal change—in regulatory standards, and in payments for harm caused—it is demonstrated that the optimal new regulatory standard is less than the conventionally efficient standard, and that the optimal new payment for harm is less than the harm.

JEL codes: H8, K10, K20

1. Introduction

One of the major categories of risk facing individuals and firms is that the law may change. Homeowners, for example, confront the prospect that the mortgage interest tax deduction will be eliminated; firms bear the risk that new regulations concerning workplace safety or environmental harm will be enacted; and fishermen live with the possibility that a fishery will be declared off-limits. Such risks of legal change are inevitable, as they are a natural byproduct of economic, technological, and political uncertainties.

A hallmark of the risk of legal change is that it is largely uninsurable. As is described in Section 2.1, insurance against legal change is for the most part nonexistent as an empirical matter. In contrast, insurance against property losses, medical needs, and all manner of other adverse events is widespread.

Why is insurance coverage against legal change mainly unavailable? I suggest in Section 2.2 that the root of the explanation lies in the correlated nature of the losses usually associated with changes in legal rules. If, for instance, the use of a fishery were barred, all fishermen earning income from it would simultaneously suffer losses. Hence, insurers could face substantial risks were they to sell coverage against changes in the law. However, as I also discuss, insurers should be able to alleviate these risks in various ways, implying that a fully satisfactory explanation of reality involves factors in addition to the correlation of losses.

In any event, because insurance against legal change is generally unavailable in fact, I assume that such coverage cannot be purchased in a model of uncertainty about the law that I analyze in Section 3. Under that assumption, as well as the assumption that parties subject to the law are risk averse, the principal analytical point of the article is developed: legal change should be attenuated, that is, legal change should be less than would be conventionally efficient (efficient in a risk-neutral world—efficient according to the customary cost-benefit calculus).

---

* Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School. I thank Louis Kaplow, Dean Lueck, A. Mitchell Polinsky, Kathryn Spier, Michael Trebilcock, and Abraham Wickelgren 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, Divya Subrahmanyam, 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.

To amplify, in the model, there is a risk that an activity will be discovered to be harmful by the state. In that case, a law addressing the harmful activity will be adopted; thus there is a risk of legal change. If the activity is found to be harmful, individuals are able to reduce the likelihood of harm by exercising costly care (effort or investment in safety).

One type of legal change that is considered is the adoption of a regulation, under which the state would mandate a level of care. (Regulation may also be interpreted as effective control of care by the courts through the application of the negligence rule.) The uninsurable risk borne by parties involved in the activity is that, if it is revealed to be harmful, a regulation will be employed and parties will have to expend resources to meet the required level of care. For example, gasoline service station owners might have to replace their gasoline holding tanks with different tanks that are less likely to leak in order to satisfy new environmental regulations. In the presence of such risks of bearing regulatory compliance costs, it is shown that the optimal regulatory standard is less than the conventionally efficient standard.

The rationale for this basic result is, in essence, that if the stipulated standard of care were equal to the conventionally efficient level, a marginal relaxation of the standard would leave expected social costs essentially unchanged.[1] But the slight 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 type of legal change that is analyzed is the adoption of strict liability, requiring a payment of damages to victims of harm. (Strict liability is effectively equivalent in the model to fines or corrective taxes.) Here the risk that individuals bear concerns damages[2] as well as costs 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 damages is less than the harm.[3] The ground for this conclusion is similar to that for the conclusion about the desirable regulatory standard of care.

In Section 4, I make several concluding comments about the model, concerning its generality and importance, the attenuation of legal change in reality (such as through delayed implementation of laws and grandfathering), and a policy of government compensation for legal change.

Before proceeding, let me comment on prior writing on which I build. The most closely related articles to this one are by Blume and Rubinfeld (1984), who address governmental takings, and by Masur and Nash (2010), who examine legal change generally. In both of these articles, the authors observe that insurance against the legal risks that they consider does not exist, and they propose that the provision of compensation by the government might be desirable as a substitute. I consider a different policy response to the absence of insurance, namely,

[1] At the conventionally efficient level of care, the marginal gain from the cost savings due to a reduction in the level of care must equal the marginal loss from the resulting increase in expected harm.

[2] More precisely, the risk is of having to pay expected damages. In particular, I assume in the model that, if strict liability is adopted, individuals can then purchase liability insurance coverage at actuarially fair rates. Therefore, individuals do not turn out to bear a risk of having to pay actual damages. However, individuals still face the risk of having to pay the insurance premium for the damages coverage that they will elect to purchase.

[3] 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 damages is zero for all harms sufficiently low; see Proposition 3.

attenuation of legal change.[4] Moreover, I offer an alternative primary explanation for the absence of insurance coverage—correlation of the losses caused by legal change.[5]

The foregoing articles are part of a larger, economically-oriented literature on legal change. This literature began with Graetz (1977) and Kaplow (1986), and continued with Logue (1993), Levmore (1999), Shaviro (2000), and Nash and Revesz (2007), among others.[6] Somewhat surprisingly, with the exception of the two articles noted in the preceding paragraph, none of the writing on legal change recognizes the unavailability of insurance against it.[7] As a partial consequence, Graetz (1977) and Kaplow (1986) come to the judgment that legal change should be guided by conventional efficiency criteria and that compensation for legal change ordinarily should not be provided. Although other writers sometimes differ in their conclusions, that is not because of the absence of insurance coverage for legal change.

Also of relevance is an earlier article of mine, Shavell (1982), which begins with an analysis of optimal liability rules in the absence of insurance. I show there that optimal damages under strict liability are less than harm, which is similar to the result here when there is a risk that a strict liability rule will be adopted. But my earlier article also concludes that the optimal standard of care equals the conventionally efficient level, which is different from the result demonstrated here when there is a risk that a new standard of care will be employed.[8]

2. The Absence of Insurance against Legal Change

2.1 The fact of the absence of insurance

Insurance against legal change is not generally sold in 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 market do not include legal change. For example, insurers such as Allstate, Geico, and Travelers list as major categories of coverage automobile, property, life, health, homeowners, liability, and a number of others, but none mention legal change as a type of coverage.[9] Second, reference works and textbooks on insurance do not describe legal change as a distinct risk for

[4] Neither Blume and Rubinfeld (1984) nor Masur and Nash (2010) analyze the policy of relaxing regulatory standards. Instead, as just stated, they focus on the policy of government compensation for legal change. Similarly, other writing about legal change (see the next paragraph) also mainly considers the policy of government compensation for legal change, not the policy of reducing regulatory standards.

[5] Blume and Rubinfeld (1984) emphasize moral hazard as the main explanation, and Masur and Nash (2007) focus on the inability of insurers to price coverage. I discuss their views in Section 2.2.

[6] See also Quinn and Trebilcock (1982), Kaplow (1992), Logue (1996), Georgakopolous (1997), Levmore (1998), and Shavell (2008).

[7] Notably, neither Graetz (1977) nor Kaplow (1986) addresses the question whether insurance against legal change exists. Kaplow, however, considers the theoretical possibility of this type of insurance (at pp. 536-550) and argues that there should be no significant problem with its efficient supply by the market.

[8] See also notes 36, 39, and 53 below on the analysis in my prior article.

[9] For an overview of Allstate's insurance products, see www.allstate.com/about/product-overview.aspx; for Geico, see www.geico.com and the pull-down menu "type of insurance"; for Travelers, see www.travelers.com.

which coverage is available.[10] Third, knowledgeable individuals whom I have contacted in the insurance industry and in academia concur that insurance coverage against legal change is chiefly unavailable.[11]

Thus, if we ask whether the risks due to legal change that I mentioned in the Introduction are insurable, the answer is no. It does not seem possible for homeowners to purchase insurance coverage against elimination of the mortgage interest tax deduction, for gasoline service station owners to arrange coverage for the expense of having to install new gasoline holding tanks or for businesses generally to secure coverage against the cost of meeting new regulatory requirements, or for fishermen to obtain coverage against bans on fishing.[12]

There is, however, an implicit exception to the absence of coverage against legal change of note. Namely, significant protection against modifications in liability rules is provided by standard liability insurance policies. Although the principal role of liability insurance is to cover insureds against the risk of liability resulting from the application of existing laws, liability insurance also shields insureds against changes in liability rules as long as the changes concern the insured categories of liability.[13] Hence, for instance, physicians and other professionals would likely be covered under their malpractice policies against expansions in their exposure to liability for malpractice. 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 at the time of policy

[10] 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 an area of coverage.

[11] 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.)

[12] To my knowledge, no insurance exists against a long term ban on fishing. It might be asked, however, whether business interruption insurance 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, pp. 226-232) and Baker (2008, pp. 41, 314-321).

[13] Baker (2004) makes this point in a broad discussion of the risks of change that liability insurance covers.

renewal,[14] from excluding new liability risks at that time,[15] or from canceling coverage altogether.[16]

There also exist scattered explicit exceptions to the absence of insurance coverage against legal change, such as coverage against expropriation of corporate assets by foreign governments[17] and individually-negotiated coverage against specific changes in tax laws.[18]

Another proviso concerning the unavailability of insurance coverage against legal change is that parties can sometimes hedge against it through the use of market transactions. For instance, homeowners facing the risk of elimination of the mortgage interest tax 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 may not be practical,[19] or limited,[20] and, in any case, require a degree of sophistication that many parties do not possess.

2.2 The explanation for the absence of insurance

Subject to the qualifications discussed above, parties are unable to insure against legal change. What is the explanation for this fact? 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 tends to affect many parties at once. Thus, as I observed, if a restriction on fishing were imposed, all fishermen subject to it would suffer concurrent losses of income, and similarly if the mortgage interest tax deduction were eliminated or if a requirement that gasoline service stations install new gasoline holding tanks were adopted, all individuals with mortgages and all service station owners would simultaneously sustain losses.

---

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

[15] See, for example, Abraham (2010, pp. 570-575) and Harrington and Niehaus (2004, pp. 616-617), noting the greater use of exclusions for pollution-related liability due to a broadening of the legal grounds for such liability.

[16] See, for example, Church (2005), discussing the cancellation of a wide array of types of liability insurance policies on account of greater liability risks.

[17] This coverage is offered by the Overseas Private Investment Corporation (OPIC); see http://www.opic.gov/what-we-offer/political-risk-insurance.

[18] Personal communication with David Bassi of Plymouth Rock Assurance Company.

[19] How would gasoline service station owners hedge against the possibility that they would have to install new gasoline holding tanks?

[20] The ability to sell short shares in home building companies seems circumscribed because the aggregate capitalization of these companies is probably 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.

It follows that if an insurer were to sell coverage against legal change, the insurer would tend to bear a substantial risk and would need to hold large reserves. Suppose that an insurer writes $1 billion of coverage against the risk of a change in the law. Then the insurer must maintain reserves fully equal to $1 billion in order to honor its contracts with policyholders—for regardless of how low the probability of claims is, such as only 1%, the triggering of its coverage responsibilities would occur simultaneously for its insureds.[21] 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 fulfill its contracts with virtual certainty.[22]

Such a 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.[23] Holding large reserves, like $1 billion, is much more expensive than holding modest reserves, like $12 million, and thus could dramatically increase insurance premiums.[24] Moreover, the risk of losing large reserves could further raise premiums if insurers are considered risk averse. Hence, the premiums that insurers would be likely to charge for provision of coverage when losses are correlated could well exceed insureds' willingness to pay.

Correlation of losses therefore seems to provide a plausible candidate explanation for the general absence of insurance coverage against changes in legal rules.

It should also be noted that the correlation-of-losses hypothesis is reconcilable with the fact that liability insurance policies furnish implicit coverage against changes in liability rules. In particular, a change in a liability rule does *not* lead to simultaneous losses for the holders of liability insurance coverage. When a liability rule changes, typically only a small minority of insureds will make claims as a consequence, for only those insureds who turn out to cause harm and are found liable as a result of the modification of the liability rule would add to the claims the insurer must pay.

---

[21] An insurer might choose to hold reserves of $1 billion in order to prevent bankruptcy and to give insureds confidence in their insurance. Another possibility is that an insurer would be forced to hold the reserves by an insurance regulator. A different possibility is that an insurer would purchase reinsurance to cover its $1 billion risk. In that case, though, the reinsurers might need to hold reserves of $1 billion in order to be able to honor their policies.

[22] Since the assumption is that the insurer writes $1 billion of coverage 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, $[.01(\$25{,}000 - \$250)^2 + .99(-\$250)^2]^{1/2}$), implying that the standard deviation of the average loss among the 40,000 insured individuals is only $\$2{,}487/(40{,}000)^{1/2} = \$12.44$. 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.

[23] See, for example, Harrington and Niehaus (2004, p. 182) and Vaughan and Vaughan (2008, p. 43).

[24] Suppose that the financing cost of maintaining liquid reserves is 4% a year. Then to hold reserves of $1 billion would cost an insurer $40 million a year. This implies that, if the insurer is risk-neutral, its premium revenue must be five times its expected losses of $10 million a year, because its premium revenue must support its financing cost and its expected losses. Such a high ratio of premiums to expected losses could discourage purchase of insurance coverage, that is, eliminate the market for coverage. In contrast, for an insurer to hold reserves of only $12 million would cost only $480,000. This implies that, if the insurer is risk-neutral, its premium revenue must be only $12.48 million, which is only1.248 times the expected losses of $10 million a year.

However, correlation of losses needs to be supplemented with other factors to satisfactorily explain the observed absence of insurance coverage against legal change. Among the reasons are that many risks of legal change affect relatively small numbers of parties,[25] that legal risks can be set off against each other,[26] and that legal risks are often uncorrelated with other categories of risk, including natural disasters and macroeconomic events. These points suggest that reinsurers and the capital markets could find risk-sharing with insureds and/or the insurance industry attractive and thus qualify the significance of correlation of losses as an explanation for the absence of coverage against legal change.[27]

Finally, let me consider several possible explanations for the unavailability of insurance that are different from correlation of losses. One standard reason for the nonexistence of insurance is adverse selection. Namely, when an insurer sells a type of coverage, say theft insurance, high-risk individuals will be more likely to buy it than low-risk individuals, implying that the price of coverage (which must reflect the average risk experienced by the insurer) will become relatively unattractive for low-risk individuals and could lead some of them not to continue to buy coverage. This could engender an unraveling process in which more low-risk individuals among the remaining insureds decide not to continue to buy theft insurance, ultimately vitiating the market for that coverage. Crucially, for adverse selection to occur, insurers must be unable to identify who are high-risk individuals, for otherwise the insurers could and would charge them appropriately high premiums.

The phenomenon of adverse selection appears to be of little importance in our context. On one hand, adverse selection could not possibly occur if insurance were offered against a specific change in the law (such as a ban on the use of the cod fishery off the coast of Maine), for then the risk would by definition be the same for all policy holders (namely, the probability of the named ban on the cod fishery)—in other words, there could not be any high-risk buyers of the coverage. On the other hand, adverse selection would be a logical possibility if insurance were offered against a category of changes in the law (such as any kind of restriction on any type of fishing on the East Coast), for then there could be high-risk buyers of coverage, that is, individuals who know that a particular legal change within the covered category is likely. But such adverse selection seems implausible, because it could occur only if a significant number of individuals know more about the likelihood of a change in the law than do insurers.

A second standard reason for the nonexistence of insurance is moral hazard, which arises when a person who purchases insurance behaves in such a way as to increase the risk of loss. This would result in an elevated premium rate and could discourage the purchase of insurance. If, for example, a person purchases insurance against loss of his cellphone, he would probably be less careful with it than otherwise; he might lose his cellphone with a fairly high probability, say

[25] 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.

[26] Consider the two legal risks mentioned in the previous note.

[27] A number of economists have asked me whether the observations of this paragraph imply that the correlation problem of insurers should be largely overcome, not merely qualified. But important evidence to the contrary is the paucity of insurance coverage against (and high premiums charged for) catastrophic risks, such as earthquakes, floods, and hurricanes—even though these correlated risks should in theory also be substantially alleviated through reinsurance and the participation of the capital markets. On the puzzle represented by the shortage of coverage against catastrophic risks and possible explanations for it, see especially Froot (2001), who emphasizes in part 5 various problems with the supply of reinsurance.

50 percent, over the course of a year, rather than with only a low probability, say 5 percent, if his cellphone is not insured. Hence, the person might well prefer not to be insured and bear a 5 percent risk of loss of his cellphone instead of paying an annual premium of 50 percent of the value of his cellphone.

Blume and Rubinfeld (1984, pp. 593-597) suggest that moral hazard could explain why insurance against takings does not exist. Their thought is that, if a person were to own takings insurance, he would be less likely to resist a taking (such as at a hearing before a zoning board). Their view seems sensible in the context of takings.[28]

But the moral hazard argument does not appear to extend from takings to the general setting of changes in the law. For instance, would a single fisherman's purchase of insurance against a ban on the use of a fishery be likely to measurably affect the probability of such a ban? I assume that the answer is no—because impeding a ban would ordinarily be a substantial undertaking and require the assent of many.[29] Hence, moral hazard should not usually operate to reduce an individual fisherman's motive to buy insurance coverage against a ban. This point must be distinguished from a different one: If a mass of fishermen were to purchase coverage against a ban on the use of a fishery, then the probability of a ban could well rise, since their collective incentive to resist a ban would be dulled. That phenomenon, however, would not necessarily depress the demand for insurance coverage—it could easily raise the demand for coverage—and thus would not constitute a clear reason for coverage not to be offered.[30]

An additional reason for the nonexistence of insurance against legal change is offered by Masur and Nash (2010, pp. 421-426), namely, that it would be difficult for insurers to price such coverage because of uncertainty surrounding the probability of legal change. However, their hypothesis does not seem consistent with the theory of insurance or with insurance practice. With regard to theory, an insurer should in principle be indifferent between pricing and selling coverage against a change in the law that will occur with a probability of 50 percent and pricing and selling coverage against a change in the law that will occur either with a probability of 25 percent or with a probability 75 percent, each with equal likelihood.[31] (This is not to deny a different point—that an insurer would decidedly prefer to sell coverage against independent and

[28] Still, other reasons for the absence of takings insurance coverage are also plausible. Notably, the demand for coverage may be low because the government often pays compensation for takings (supposedly always for physical takings of property, and sometimes for regulatory takings).

[29] In contrast, a single individual faces a much less daunting task if he is seeking only to prevent the application of takings law to his particular parcel of land.

[30] After a mass of fishermen buy coverage and the probability of a ban rises, the demand for insurance coverage could either increase or decline. In particular, it is well known that the demand for insurance is a function of the probability p of loss, rising from zero when p is 0 to a maximum at some probability p^* less than 1, and then falling to zero when p is 1. (In particular, if $x > 0$ is the possible loss, U a person's utility function, and y his wealth, then the expected utility value of insurance coverage (which reflects the strength of demand for coverage), if sold at actuarially fair rates, is given by $U(y - px) - [(1 - p)U(y) + pU(y - x)]$. This function equals 0 at $p = 0$ and $p = 1$, achieves a maximum at p^* determined by $U(y) - U(y - x) = xU'(y - px)$, and is concave.) Thus, for example, if the probability of a ban is less than p^* before insurance is sold and increases to a probability that is still less than p^* after insurance is sold, the demand for insurance will increase; and if the probability of a ban exceeds p^* before insurance is sold, then after insurance is sold the demand for insurance will decrease.

[31] In both cases, the insurer would face a 50 percent risk of a change in the law, and thus a 50 percent risk that all insureds would simultaneously make claims for loss.

identical risks of known magnitude, such as 50 percent, than to sell coverage against independent and identical risks of unknown magnitude.[32]) With regard to insurance practice, insurers routinely sell negotiated policies against a wide range of risks that would be described as subjective (such as the risk of the destruction of the World Trade Center or the risk of harm to actress Betty Grable's legs). Moreover, if the subjective nature of the risks of legal change were a bar to its sale by insurers, why would liability insurers sell coverage that includes protection against legal change?

To summarize and conclude the discussion, we have seen in Section 2.1 that, as a matter of observation, insurance coverage against legal change is mainly unavailable. It has also been suggested in Section 2.2 that the correlation of losses due to legal change provides a partial explanation for the absence of coverage and that several major alternative explanations for it seem problematic. Nevertheless, the correlation hypothesis is not entirely satisfying. In any case, given that insurance against legal change is largely unavailable in reality, I will make the simplifying assumption in the model below that insurance against legal change does not exist.

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 will learn at a future date whether the activity of injurers is potentially harmful or is instead harmless. If the activity is determined to be potentially harmful, the state will announce a legal rule to address the danger. Thus, the risk of legal change is the risk that the state discovers that the activity is potentially harmful and as a consequence enacts a legal rule to control the danger. After the announcement of the rule, injurers choose a level of care to reduce the probability of harm.[33] 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; $0 < p(x) < 1$; $p'(x) < 0$; $p''(x) > 0$; and

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

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

[32] To illustrate, suppose that an insurer sells coverage against many identical, independent risks of 50% of loss of $100 each. Then, by the law of large numbers, if the insurer charges a premium of slightly over $50, it will be virtually certain to cover the claims it will have to pay (see also note 22). Now suppose that the insurer is uncertain about the common risk of loss—that it is either 25% or 75%, each with equal probability. Then even though 50% is the expected risk, it would be folly for the insurer to charge a premium in the neighborhood of $50, for then if the common risk turns out to be 75%, the insurer would go bankrupt. In fact, the insurer would have to charge a premium slightly above $75 to avoid a 50% chance of bankruptcy. One way to express these points is that in the classic context of insurance, where the insured risks are independent and identical, uncertainty about the common risk generates a species of correlation problem. See, for example, Harrington and Niehaus (2004, p. 182).

[33] Although this model captures in a stylized way the basic elements of many legal changes, it does not directly describe all examples of legal change, such as the possible elimination of the mortgage interest tax deduction. But some such examples can be accommodated with minor modifications of the model (see note 47 on the mortgage interest tax deduction), and in any case, it does not seem that the main qualitative conclusion depends on the particulars of the model (as is suggested by note 41).

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

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.[34] Since the case in which x^* is zero is uninteresting, it will be supposed that x^* is positive, or equivalently, that

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

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) $p'(0)h < -1$,

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 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 nature of most 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.[35] 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:

---

[34] See, for example, Landes and Posner (1987) and Shavell (1987).

[35] 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.) It also does not make sense to employ the sum of expected utilities as the optimality criterion, for although that goal would be promoted by bettering the allocation of risk, it would also be furthered by redistribution from wealthy to poor. (Hence, if some policy advances the sum of expected utilities, the explanation could be that the policy favorably redistributes wealth rather than that it improves the allocation of risk.)

(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.

I now consider Pareto optimal legal rules when the rules concern regulation of behavior and then when they concern payments for harm.[36]

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 (through public enforcement or private, under the negligence rule) would be distracting for our purposes.

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

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

where r is a reference level of expected wealth. The Pareto optimal standard x_s maximizes (6) subject to (7).[37] From (7), we have

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

so that the problem at issue reduces to maximizing

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

over x_s. Denote the Pareto optimal standard x_s that maximizes (9) by x_s*.[38]

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*.

---

[36] In Shavell (1982), I consider Pareto optimal legal rules in an accident model with possibly risk-averse parties, but with the crucial difference that there is no uncertainty about legal change—whether a legal rule will apply. See notes 39 and 53 below for further comments on my prior article.

[37] That is, maximization of (4) subject to (5) is maximization of (6) subject to (7) in the present context.

[38] Since $EU''(x_s) < 0$ will be shown in (11) below, x_s* is unique.

The claim in this remark holds 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 $x*$[39] 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.[40] Thus, if the injurer were risk neutral, the 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.[41]

Proof. First, observe that

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

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

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

since each term is negative.

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

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

---

[39] In Shavell (1982), I considered the Pareto optimal standard of care assuming, unlike here, that possibly risk-averse injurers know ex ante that their activity of injurers is subject to control. Then the Pareto optimal standard equals the conventionally efficient level $x*$; see Proposition 3(a) in my earlier article.

[40] 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 $-qp'(x*)h = q$.

[41] Note that this logic explaining the key result that $x_s* < x*$ does not depend on the fact that, in the model, x_s affects other parties by influencing expected harm, $p(x_s)h$. It is readily verified that $x_s* < x*$ would still be true if x_s were to affect other parties via any positive cost function $f(x_s)$ that is decreasing in x_s, in which case (7) would be replaced by $v + t - f(x_s) = r$ and $x*$ would be the x minimizing $x + f(x)$.

$$= -q(p'(0)h + 1)U'(w - qh(0)) > 0$$

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) $$\begin{aligned} EU'(x^*) &= -qp'(x^*)h(1-q)U'(w - qp(x^*)h) - (qp'(x^*)h + 1)qU'(w - qp(x^*)h - x^*) \\ &< -qp'(x^*)h(1-q)U'(w - qp(x^*)h - x^*) - (qp'(x^*)h + 1)qU'(w - qp(x^*) - x^*) \\ &= [-qp'(x^*)h(1-q) - (qp'(x^*)h + 1)q]U'(w - qp(x^*)h - x^*) \\ &= -q(p'(x^*)h + 1)U'(w - qp(x^*)h - x^*) = 0, \end{aligned}$$

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

The first-order condition determining x_s^* is

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

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.[42] □

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

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

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.[43] The Pareto optimal standard is therefore determined by maximizing (6) subject to

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

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

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.

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 (A11) 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.

---

[42] I will not comment on first-order conditions in later propositions because their interpretations are similar.

[43] 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 essential nature of the conclusion I draw in the next proposition would not change.

(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. See the Appendix.

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.[44] Under strict liability, injurers must pay damages to victims if harm occurs. Let

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

In our legal system, d is generally intended to equal h, but here we will determine the Pareto optimal d, which need not equal h. 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 $p(x)d$.[45] (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 risk that the state will employ the rule of strict liability. In any event, the assumption that liability insurance can be purchased is not essential.[46]) Hence, the injurer's utility will be

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

Under corrective taxes, injurers are required to make a payment of $p(x)d$ if the state learns that the activity is harmful.[47] 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 (17).

---

[44] A fine paid to the state would also be equivalent to damages paid to victims given the assumptions in the model.

[45] This presumes that liability insurers can observe x and base the premium on it; see, for example, Shavell (1987, chapter 8).

[46] 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.

[47] It would be straightforward to consider a variation of the model in which, in the initial situation before uncertainty is resolved, a subsidy (a negative corrective tax) is enjoyed by the individuals with utility function U because of the existence of a positive externality generated by their choice of x. Then, it would be discovered with probability q that the positive externality does not exist. Hence, the Pareto optimal subsidy would fall to zero if the parties who had enjoyed the subsidy were risk-neutral, but the subsidy would not fall to zero if they were risk averse. This variation of the model can be viewed as describing the possible elimination of the mortgage interest tax deduction, since it is a form of subsidy that could be interpreted as justified by a positive externality associated with home ownership.

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

The level of care that an injurer chooses maximizes (17),[48] so that it minimizes

(18) $x + p(x)d$.

Since (18) is strictly convex in x, there exists a unique x minimizing it, which will be denoted $x(d)$. Further, since the derivative of (18) 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

(19) $1 + p'(x)d = 0$.

From implicit differentiation of (19), we obtain that $p''(x)x'(d)d + p'(x) = 0$, implying that

(20) $x'(d) = -p'(x)/(p''(x)d) > 0$.

If victims are risk neutral, the expected wealth of a victim will be $v + t - qp(x)(h - d)$,[49] since victims bear $h - d$ of their losses.[50] Thus, we must have

(21) $v + t - qp(x)(h - d) = r$,

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

(22) $t = r - v + qp(x)(h - d)$.

The Pareto optimal level of damages d for risk-averse injurers and risk neutral victims thus maximizes (17) subject to (22). Hence, d^* is determined by maximization of

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

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, (23) reduces to

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

To maximize (24), $[p(x(d))h + x(d)]$ 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 (32) below.

[48] 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.

[49] Here and below, I will sometimes write x instead of $x(d)$ for notational convenience.

[50] 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.

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.[51] 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 the 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;[52] that is, if the 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.[53]

Proof. I first show that $EU(d)$ is strictly decreasing in d for d in $[0, -1/p'(0)]$. From (23), we have

$$(25)\quad \begin{aligned} EU'(d) = {} & (1-q)[-qp'(x)x'(d)(h-d) + qp(x)]U'(u-r+v-qp(x(d))(h-d)) \\ & + q[-qp'(x)x'(d)(h-d) + qp(x) - x'(d) - p'(x)x'(d)d - p(x)] \\ & \times U'(u-r+v-qp(x)(h-d)) - x(d) - p(x)d). \end{aligned}$$

Now for d in $[0, -1/p'(0)]$, recall that $x(d) = 0$. Hence, $x'(d) = 0$ in the interval. Thus, in that interval (25) reduces to

$$(26)\quad \begin{aligned} EU'(d) = {} & (1-q)[qp(0)]U'(u-r+v-qp(0)(h-d)) \\ & + q[qp(0) - p(0)]U'(u-r+v-qp(0)(h-d)) - p(0)d). \end{aligned}$$

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

$$(27)\quad \begin{aligned} EU'(d) < {} & (1-q)[qp(0)]U'(u-r+v-qp(0)(h-d) - p(0)d) \\ & + q[qp(0) - p(0)]U'(u-r+v-qp(0)(h-d) - p(0)d) = 0. \end{aligned}$$

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 (25) we see that

$$(28)\quad \begin{aligned} EU'(d) < {} & \{(1-q)[-qp'(x)x'(d)(h-d) + qp(x)] \\ & + q[-qp'(x)x'(d)(h-d) + qp(x) - x'(d) - p'(x)x'(d)d - p(x)]\} \\ & \times U'(u-r+v-qp(x)(h-d)) - x(d) - p(x)d) \end{aligned}$$

---

[51] 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.

[52] Since $t = r - v + qp(x(d))(h-d)$, the increase in t from a marginal reduction in d is $-qp'(x)x'(d)(h-d) + qp(x)$, 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.

[53] In Shavell (1982), I considered the Pareto optimal level of damages assuming, unlike here, that possibly risk-averse injurers know ex ante that their activity is subject to strict liability. Then, when injurers can purchase liability insurance, the Pareto optimal level of damages d^* equals the harm h; see Proposition 4 in my earlier article.

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

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 (28) 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 $m(h)$ = maximum of $EU(d, h)$ 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 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 (25), making use of the fact that $x'(d)(1 + p'(x)d) = 0$ since $x(d)$ satisfies $1 + p'(x)d = 0$, to obtain

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

We want to show that (29) is negative for $d \geq h$. Since $-qp'(x)x'(d)(h-d) \leq 0$ for $d \geq h$, it suffices to show that

(30) $$(1-q)qp(x)U'(u - r + v - qp(x(d))(h-d))$$
$$+\, q[qp(x) - p(\mathrm{x})]U'(u - r + v - qp(\mathrm{x})(h-d)) - x(d) - p(x)d) < 0$$

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

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

The first-order condition determining d^* is, from (29),

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

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

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

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 (17) subject to

(34) $$(1-q)V(v + t) + qV(v + t - p(x(d))(h-d)) = r,$$

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 (A27) 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.[54]

(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. See the Appendix.

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 is announced. If other types of action were considered, notably, decisions about levels of care before a legal change, or decisions about levels of activity before or after a legal change, nothing essential would be altered. There would then be different dimensions of conventional inefficiency of behavior engendered by the attenuation of legal change, but some attenuation from efficiency would still generally be desirable to relieve risk-bearing.[55]

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 not alter the main conclusions.

A third simplifying assumption was that there was no insurance available against legal change. If partial insurance were available, risk-averse parties would still bear some risk, so that attenuation of the legal change might be desirable.

(b) *The desirable extent of attenuation depends on the degree of risk aversion and risk.* The optimal level of attenuation of legal change is a function of the degree of risk aversion of parties and of the magnitude of the risks of losses that they bear. Thus, for individuals subject to substantial legal risks, such as homeowners with large mortgages confronting the potential elimination of their interest deductions, significant attenuation would presumably be desirable, whereas for publicly-held firms facing the possibility of a modest change in a workplace safety rule, essentially no attenuation would be appropriate.

[54] In Proposition 3, this logic did not apply since victims were risk neutral.

[55] For example, it should be clear to a reader who reviews Note (b) to Proposition 1 and note 41 that the proof that the optimal standard x_s^* is less than the conventionally efficient level x^* would be essentially unaltered if x were chosen before the state learns whether the activity is harmful. (Both x_s^* and x^* would be lower under that assumption, but $x_s^* < x^*$ would still hold.)

(c) *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 convince regulators to lessen the rigor of the regulation. Moreover, legal change is frequently delayed or implemented in phases. A 10 percent increase in a tax rate might be achieved through a 5 percent increase in one year and another 5 percent increase in a second year. Also of relevance is grandfathering, under which parties who are already engaged in an activity governed by a legal rule are exempted from having to satisfy a legal change bearing on it. Grandfathering is obviously a form of attenuation, as it arrests legal change for the parties engaged in an activity, and grandfathering possesses the beneficial feature that these parties may be least able to bear the risk of change (parties who are only contemplating entry into the activity have the option to delay doing so or not to enter at all). In sum, legal change seems to display a broadly attenuated character in reality, which is consistent with the analysis of this article.[56]

(d) *Attenuation of legal change versus government compensation.* Although it was shown in this article that attenuation of legal change is a desirable policy given the unavailability of insurance coverage against modifications of the law, other policies were not considered. One such policy is government compensation for the losses caused by legal change, and I restrict my remarks here to it.[57]

Consider the risk of losses due to adoption of a new regulation. Then a policy of government compensation might allow achievement of an ideal outcome. In particular, full compensation could be granted for the cost of meeting a new regulation (such as the cost to a gasoline service station of installing new, mandated holding tanks), so that no risk would be borne by parties subject to the regulation. Hence, there would be no reason for society to lower the required level of care from the conventionally efficient level (no reason to lower the required quality of the holding tanks from the efficient level), and the best possible outcome would apparently be obtained.[58]

Next consider the risk of losses due to the use of a legal rule requiring a payment. Here, a policy of government compensation would be equivalent to attenuation of the payment. For example, suppose that a pollution tax of $10 per pound of pollutant might be enacted and that, if it is, the government would provide $3 of compensation per pound of the taxed pollutant. This program would clearly lead to the same result as a tax of $7 per pound unaccompanied by any

---

[56] That legal change is often observed to be attenuated may also be explained in part by a rationale that is independent of risk aversion: to avoid wasting resources already committed to comply with the law (if a gasoline service station recently installed a holding tank that met then-relevant regulations, it could be wasteful to force the service station to replace that tank with a slightly more efficient one). This point is developed in Shavell (2008). The attenuation of legal change may also be ascribed to the self-interest and political power of parties who would be adversely affected by a legal change (whether or not attenuation has an economic rationale).

[57] As I noted in the introduction, the economically-oriented literature on legal change focuses on the policy of government compensation.

[58] I say “apparently” because the argument just given elides the following question: Does the cause of failure of the private market to supply insurance coverage against legal change also bear on the government’s ability to provide compensation? Notably, if the cause of the failure of the insurance market to supply coverage is the correlation of losses caused by a legal change, would the correlation of these losses also constitute a problem for a government compensation program (which must ultimately be financed by taxpayers)?

compensation. Hence, since I considered all possible levels of attenuation of a corrective tax, I implicitly considered the policy of government compensation for losses due to a corrective tax.

A more expansive analysis of legal change and aversion to risk than that of this article would include government provision of compensation or government provision of insurance coverage, as an alternative to, or complement of, attenuated legal change.

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.

Georgakopoulos, Nicholas L. 1997. "Predictability and Legal Evolution," 17 *International Review of Law and Economics* 475-489.

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. 1982. "On Liability and Insurance," 13 *Bell Journal of Economics* 120-132.

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.

Appendix

Proof of Proposition 2. 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

(A1) $t'(x_s) = qp'(x_s)hz(x_s) < 0,$

where

(A2) $z(x_s) = V'(v + t(x_s) - p(x_s)h)/[(1 - q)V'(v + t(x_s)) + qV'(v + t(x_s) - p(x_s)h)].$

Note that

(A3) $z(x_s) > 1.$

Differentiation of (A1) gives

(A4) $t''(x_s) = qp''(x_s)hz(x_s) + qp'(x_s)hz'(x_s).$

Since $z'(x_s) < 0$,[59] (A4) implies that

(A5) $t''(x_s) > 0.$

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

(A6) $EU(x_s) = (1 - q)U(u - t(x_s)) + qU(u - t(x_s) - x_s)$

over x_s. We have

(A7) $EU'(x_s) = -(1 - q)t'(x_s)U'(u - t(x_s)) - (t'(x_s) + 1)qU'(u - t(x_s) - x_s).$

Hence,

(A8) $$\begin{aligned} EU'(0) &= -(1 - q)t'(0)U'(u - t(0)) - (t'(0) + 1)qU'(u - t(0)) \\ &= -(t'(0) + q)U'(u - t(0)). \end{aligned}$$

But

(A9) $t'(0) = qp'(0)hz(0) < qp'(0)h < -q,$

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

(A10) $$\begin{aligned} EU(x_s) &= (1 - q)U(u - t(x_s)) + qU(u - t(x_s) - x_s) \\ &= (1 - q)(u - t(x_s)) + q(u - t(x_s) - x_s) = u - (t(x_s) + qx_s) \end{aligned}$$

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

(A11) $t'(x_s) = -q.$

Given (A11) 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 (A7), we know that x_s^* is determined by

(A12) $-(1 - q)t'(x_s)U'(u - t(x_s)) - (t'(x_s) + 1)qU'(u - t(x_s) - x_s) = 0,$

which implies that at x_s^*,

(A13) $$\begin{aligned} t'(x_s) &= -qU'(u - t(x_s) - x_s)/[(1 - q)U'(u - t(x_s)) + qU'(u - t(x_s) - x_s)] \\ &< -qU'(u - t(x_s) - x_s)/[(1 - q)U'(u - t(x_s) - x_s) + qU'(u - t(x_s) - x_s)] = -q, \end{aligned}$$

establishing the claim.

Last, let us show that $x^* < x^{**}$. From (A11) and (A1), we have

(A14) $qp'(x^{**})hz(x^{**}) = -q,$

which implies that

(A15) $p'(x^{**})h = -1/z(x^{**}) > -1$

since, by (A3), $z(x^{**}) > 1$. Because $p'(x^*)h = -1$ and $p''(x) > 0$, (A15) implies

[59] 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.

that $x^{**} > x^*$.

Proof of Proposition 4. Equation (34) determines $t(d)$, and differentiating (34) with respect to d gives

$$\text{(A16)}\quad (1-q)t'(d)V'(v+t(d)) + q[t'(d) - p'(x)x'(d)(h-d) + p(x(d))]V'(v+t(d)-p(x(d))(h-d)) = 0.$$

Hence,

$$\text{(A17)}\quad t'(d) = q[p'(x)x'(d)(h-d) - p(x(d))] \times V'(v+t(d)-p(x(d))(h-d))/[(1-q)V'(v+t(d)) + qV'(v+t(d)-p(x(d))(h-d))].$$

The optimal d is determined by maximizing

$$\text{(A18)}\quad EU(d) = (1-q)U(u-t(d)) + qU(u-t(d)-x(d)-p(x(d))d),$$

the derivative of which is

$$\text{(A19)}\quad EU'(d) = -(1-q)t'(d)U'(u-t(d)) - q[t'(d) + x'(d) + p'(x)x'(d)d + p(x(d))]U'(u-t(d)-x(d)-p(x(d))d).$$

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

$$\text{(A20)}\quad t'(0) = -qp(0)V'(v+t(0)-p(0)h)/[(1-q)V'(v+t(0)) + qV'(v+t(0)-p(0)h))]$$

and

$$\text{(A21)}\quad EU'(0) = -(1-q)t'(0)U'(u-t(0)) - q[t'(0)+p(0)]U'(u-t(0)) = -[t'(0)+qp(0)]U'(u-t(0)).$$

But from (A20), we have

$$\text{(A22)}\quad t'(0) + qp(0) = qp(0)\{1 - V'(v+t(0)-p(0)h)/[(1-q)V'(v+t(0)) + qV'(v+t(0)-p(0)h))]\} < 1.$$

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, (A19) reduces to

$$\text{(A23)}\quad EU'(d) = -(1-q)t'(d)U'(u-t(d)) - q[t'(d) + p(x(d))]U'(u-t(d)-x(d)-p(x(d))d) = -t'(d)[(1-q)U'(u-t(d)) + qU'(u-t(d)-x(d)-p(x(d))d)] - qp(x(d))U'(u-t(d)-x(d)-p(x(d))d).$$

To show that (A23) is negative for $d \geq h$, let us first verify that $t'(d) > -qp(x(d))$ for $d \geq h$. Using (A17), the latter inequality is equivalent to

$$\text{(A24)}\quad q[p'(x)x'(d)(h-d) - p(x(d))] \times V'(v+t(d)-p(x(d))(h-d))/[(1-q)V'(v+t(d)) + qV'(v+t(d)-p(x(d))(h-d))] \geq -qp(x(d)).$$

Since $p'(x)x'(d)(h-d) \geq 0$ when $d \geq h$, (A24) will be true if

$$\text{(A25)}\quad -qp(x(d)) \times \{V'(v+t(d)-p(x(d))(h-d))/[(1-q)V'(v+t(d)) + qV'(v+t(d)-p(x(d))(h-d))]\} \geq -qp(x(d)).$$

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 (A23) for such d that

$$\text{(A26)}\quad EU'(d) \leq qp(x(d))[(1-q)U'(u-t(d)) + qU'(u-t(d)-x(d)-p(x(d))d)] - qp(x(d))U'(u-t(d)-x(d)-p(x(d))d) = qp(x(d))\{[(1-q)U'(u-t(d)) + qU'(u-t(d)-x(d)-p(x(d))d)] - U'(u-t(d)-x(d)-p(x(d))d)\} < 0.$$

Hence, $d^* < h$.

The first-order condition determining d^* is

$$\text{(A27)}\quad -t'(d)[(1-q)U'(u-t(d)) + qU'(u-t(d)-x(d)-p(x(d))d)] = qp(x(d))U'(u-t(d)-x(d)-p(x(d))d).$$