

Insurance and Catastrophes

RICHARD ZECKHAUSER

Harvard University, Kennedy School of Government, 79 Kennedy St., Cambridge, MA 02138

Abstract

Catastrophes provide a principal justification for insurance. Traditional conceptions of catastrophes miss three critical elements. (1) Many catastrophes—the liability revolution in the United States, for example—are not bolts from the blue. Rather, they develop over many years and result from human activity. (2) Conventional, experience-based models for assessing losses often smudge the distinction, so critical for catastrophes, between probability and magnitude of loss. (3) Normal insurance contracts, with heavy copayments for small losses but little charge at the margin for large ones, perform poorly when the insured can tradeoff probability and size of loss—a phenomenon we label distribution distortion. The structures of optimal insurance contracts are assessed.

Key words: catastrophe, insurance, moral hazard, copayment, experience rating, distribution distortion

1. Introduction

Catastrophes provide the predominant conceptual model of what insurance is about. One pays premiums to secure financial protection against low-probability high-consequence events—what we normally label catastrophes. In this paper I offer a few thoughts about catastrophes that may diverge from conventional wisdom.

First, I argue that while we usually think of catastrophes as a bolt from the blue—say, an earthquake, flood, or explosion—this model does not capture many of the events that are the most consequential for insurance markets. The elusive events—the liability revolution in the United States is an example—are often predominantly the result of human activity, and they may develop over many years. Second, traditional methods for updating predictions of loss on the basis of experience may not incorporate information from catastrophic events appropriately. Third, the traditional forms of insurance contracts, which impose heavy copayments on small losses but little or nothing at the margin on large ones, may be ill suited to situations where the insured's behavior can affect the size of its losses, what I label distribution distortion hazard.

These arguments are predominantly methodological. However, to appeal to insurance practitioners, the analysis is presented intuitively.

2. Catastrophes from human activities

Newspapers do not report many of the catastrophes that are most important to either individuals or insurance companies. When a family of four dies in an earthquake, they and

their relatives do not much care how many others died alongside. Similarly, for the middle-aged man who loses his job, this is a catastrophe, whether or not his loss is part of a whole-sale layoff or massive downturn. For insurers, earthquakes in Kobe or Northridge are assuredly disasters, but so too are situations where thousands of women claim to be damaged by breast implants or where buildings are hampered by asbestos, whether dangerous or not.¹ In the 1980s in the United States, we witnessed a liability explosion of catastrophic proportions; it arose not because thousands of doctors slipped their scalpels but rather because conditions in courtrooms changed. Accidents that would have won no recovery in the 1970s frequently won big awards in the 1980s.

2.1. Aggregate catastrophes

Aggregate catastrophes occur when many similarly situated people, all subject to common risks, suddenly find that they have suffered a loss, and the total losses well exceed expectations. In technical language, the common probability has shifted. Should the payoffs from insurance depend on the aggregate circumstances? In some sense this question is preposterous. Why should A get paid more or less just because B and C suffered at the same time?

At first glance, it seems irrational to give out the most funds when conditions are worst and presumably insurance funds are most valuable. However, there is a solid rationale for higher payouts when many entities suffer a common loss. Adverse selection and moral hazard are both severe concerns when writing insurance. When B and C suffer a loss at the same time as A, this is some evidence that the loss was beyond A's control. When B and C are also losers, it suggests that A neither knew a priori that his risk was high nor behaved poorly.²

For reasonable models, if adverse selection and moral hazard are serious concerns, insurance should pay more the more people suffer a common fate. For example, when a single home burns to the ground, it is quite likely that careless smoking, inadequate wiring, or some other condition under the homeowner's control was at fault. But when a whole neighborhood burns to the ground, it is more likely that an uncontrollable fire was the culprit.³

The number of people suffering a common fate is information that helps tell us how much control they had over their fate. This useful information should be employed in formulating insurance policies for catastrophes.⁴

2.2. Unpredictable risks

A traditional challenge for insurers is to determine the risk associated with various products and processes. In recent years, at least in the United States, insurers have confronted the bigger challenge of predicting how courts and legislatures will determine what products actually caused harm and what payments they will require. For example, in the Exxon-Valdez case, once the spill occurred and the immediate damages were known, there was still enormous uncertainty on how the damages would be estimated. (Uncertainties about how swiftly or completely the affected ecosystems would recover were less consequential, since that information did not substantially affect payments for damages.) In general, the

primary contemporary uncertainties on the clean-up of hazardous waste sites include what reforms, if any, will be made about retroactive liability; standards and prioritization for clean up (including any possible use of risk-benefit analysis); and the treatment of joint and several liability. Uncertainties about traditional insurance considerations, such as what materials will leak and whether they will cause injury, are not as significant.

Liability revolutions change probabilities of a common occurrence for millions of individuals. If the probability that an incident—say, one related to malpractice or product liability—rises, the costs of insurance will increase apace.

Was not the liability revolution, as opposed to, say, natural disasters, the cause of the problems of Lloyd's of London? We used to hear how Lloyd's was successful at insuring a pianist's hands or a ballerina's legs or other hard-to-estimate risks. Such efforts are mere child's play compared to estimating how a legal and judicial system will evolve.⁵ Most true catastrophes may be the work of society, with nature little or not involved.

The experience in the United States in recent years is instructive. Its impressive long-term trend in reducing accident rates is well known. For example, the overall death rate from unintentional injuries per 100,000 people per year fell from the eighties in the first two decades of this century, to the seventies through the end of World War II. It had fallen below fifty by 1975, forty by 1983, and thirty-five by 1993. Since 1975, motor vehicle death rates have fallen by 20 percent, workplace deaths by 40 percent, and home deaths by 25 percent (U.S. National Safety Council, 1994, pp. 26–27).

Yet from 1975 through 1989, the number of personal injury products liability cases commenced in federal courts increased almost sixfold. This does not simply reflect the litigation explosion; these cases grew from 2.04 to 5.74 percent of all federal civil cases. The big explanation for this growth was asbestos. Asbestos cases comprised 60 percent of products liability cases by the late 1980s. Nonasbestos liability cases increased dramatically through 1981 but then tapered off to become roughly their starting percentage of all civil cases (Viscusi, 1991, pp. 7–22).

Thus, for companies insuring the U.S. market, the megacatastrophes of 1975 to 1989 were the sleeping dog asbestos problem and the general liability explosion. Viscusi, a primary commentator on these events, identifies the principal culprits in the liability cost explosion as the emergence of strict liability for producers, hazard warnings cases (producers are sued for not alerting customers to risks), and, most important, the new phenomenon of mass toxic torts (large numbers of parties sue for being injured in a common fashion). Mass toxic torts frequently hinged on the concept of retroactive liability (Viscusi, 1991).

2.3. Massive losses and the role of government

A good reason for limiting losses in a catastrophe is that insurance companies simply can't pay. We have seen Lloyd's of London totter on the brink of insolvency. In the early 1960s, U.S. insurance companies debated whether policies should say that they would not pay off should there be a widespread nuclear war. Practically, of course, it did not matter, since the assets supporting the insurance, at least in the United States, would be significantly destroyed. But catastrophes of far lesser consequence than a nuclear holocaust could bankrupt insurance companies. For example, should Tokyo be hit by a massive earthquake, it

has been estimated that losses could easily exceed a trillion dollars. Present insurance arrangements, it would seem, would be inadequate to deal with such a situation. It should be understood, implicitly or explicitly, how resources will be divided in such circumstances.

Interestingly, after many financial and natural catastrophes, the largest payouts come from governments.⁶ Governments have the ability to tax in the future, giving them a resource base that dwarfs that of any insurer. Thus, in my country the government doled out vast amounts during the S&L crisis, and whenever California was hit by an earthquake or other plague. Hypothetically, under the Price-Anderson Act, in a nuclear catastrophe it would also pay amounts over \$600 million.

Other government and private insurance programs compensate individuals who have suffered losses that to them are catastrophic but are not part of a general catastrophe. This could include the loss of a job or the loss of one's house to fire. Interestingly, unemployment insurance is often more generous the worse are economic conditions. For example, those who live in areas where unemployment is above a specified level may get extensions of insurance, and legislatures tend to vote more generous provisions when the nation is in recession. Widespread unemployment in this context can be thought of as an aggregate catastrophe.

2.4. Economic catastrophes

Economic catastrophes—the crash of the Mexican peso at year end 1994, of the securities markets in 1987, or a collapse in confidence that brings about a recession—are assuredly created by people, and they are events that are extremely hard to estimate actuarially. Indeed, this is why we have relatively little insurance written on economic events.⁷ Many forms of insurance, however, do relate to economic events. For example, when a bank writes a mortgage on an office building, it is, in effect, insuring the buyer against a catastrophic fall in value, as banks around the world have learned to their discomfort when falling real estate prices turned them into involuntary landlords.

Given the difficulties that private insurers may have with insuring against economic events, and given that governments might even exploit such insurance if it were written, it often falls to government to insure against economic catastrophes. In the United States, at least, such instruments as unemployment insurance and bank deposit insurance are supplied by the government.

The U.S. savings and loan crisis suggests that this is not always a desirable arrangement. A system that is underpinned by politics may not be well suited to make tough decisions on pricing appropriately and denying insurance. Moreover, government insurance policies may be compelled to operate by formulas, whereas competing private insurers could take a range of factors into account.⁸ Once the government takes on an insuring role, it may be forced to regulate or preach to its insureds to prevent them from taking excessive risks.⁹

3. Catastrophes and the prediction of losses

In the summer of 1960, I worked as an actuarial student in the group insurance department at the Equitable Life Assurance Society. My assignment was a surprising one for someone

who knew even a little bit about insurance. I was to figure out—which mostly meant compute—how much it would cost if Equitable's group life insurance policies allowed their customers to write off catastrophes (that is, have them ignored in setting dividends and future premiums).

Our group insurance arrangements were fairly traditional. To some extent Equitable absorbed risk; however, as a mutual, it also paid dividends. The calculations that determined an insured's dividends gave great weight to experience; the better your experience, the more you got back. Until 1960, a customer's dividend account could be substantially in deficit—say, because there had been an explosion that took twenty lives (the example that was actually used at the time). Equitable was concerned that such clients would switch to other insurers, having little hope that Equitable would resume the dividend payments anytime soon.

The puzzling question was, why should we, the insurer, care? Doesn't the customer's bad experience reveal that it is a high-risk insured and that it should pay higher premiums? If other insurers know about the bad experience, presumably they would charge the company extra; if they didn't find out, then we would have sloughed off a loser.

Needless to say, no one thought it necessary to explain to the actuarial student the Equitable's rationale for caring, for considering the writeoff. Thirty-five years later, it seems reasonable to return to the problem of my youth to seek a logical rationale.

The central question is, in a system of adjust-premiums-as-you-go, what is the justification for not taking full consideration of the magnitude of losses when they are bunched? In other words, why might we be willing to write off catastrophes? I see three justifications: first, risk aversion swamps informational content, hence the risk-neutral party should absorb any costs, and would agree to do so if the contract were drawn before the catastrophe occurred; second, given ties between the insured and insurer, hence market power, the insurance company might lose some customers who are profitable even though they have experienced some catastrophes that they are not being fully charged for; third, the mean of losses is not a sufficient statistic. Let us evaluate these justifications for writing off catastrophic losses in turn.

First, the writeoff could be a legitimate response to the risk aversion of the insureds, particularly justified if even the highest-risk groups suffer relatively low levels of losses from catastrophes. Consider the following situation for 200-person firms. On average, ten people die in a year. The insurance includes a catastrophic component, which provides that if six or more people die in a bunch it is treated as a single death. Catastrophe-prone firms have one every 100 years; catastrophe-safe firms have one every 200 years. In such circumstances, for practical purposes, the risk-aversion advantages of writeoffs swamp risk-updating considerations as far as catastrophes are concerned. Writing off catastrophes does little harm to pricing efficiency. However, a significant excess of ordinary deaths should not be written off, assuming that some firms have worse experience on average than others.

The second argument starts from the observation that insurance markets are not really competitive in many circumstances. This lack of competition may stem from information asymmetries. For example, a potential new provider's concerns about adverse selection may make it difficult for an employer to switch to a new firm for its group insurance. Thus, insurers and their customers tend to be tied to each other. The equilibrium in such a situation is that when companies are first shopping for insurance, or can show that bad experience is not what is causing them to shop, an insurance firm will offer lower premiums

as inducements. The insurance companies count on earning higher profits during the course of a long-term relationship.

Even bad-experience firms may be able to get more favorable treatment if they are not fully distinguishable from "honest" shoppers. If the old insurer merely boosted the premium for bad firms in accord with experience (say, maintaining the same absolute profit level), these firms might leave. Thus, to preserve an old, albeit diminished, profit stream, the insurer might choose to charge less in this relationship than actuarial differentials from its other customers would suggest it should. The insured might argue, "We have taken a loss. I can place my insurance readily elsewhere, since information does not flow fully. Therefore, you and I had better share my loss."¹⁰ This can be done by writing off catastrophes, in whole or in part.

The third argument, one that is particularly appropriate for a discussion of insurance and catastrophes, is that the losses associated with catastrophic experiences may not predict future losses as strongly as would equivalent ordinary losses. (Surprisingly, even for relatively common occurrences such as auto accidents, many insurance policies update much too little and hence subsidize known high risks.)¹¹ That is, the number of deaths in a single incident does not tell us much about risk levels, since the overwhelming number of deaths are from ordinary causes, such as from heart attacks or strokes.

In an extreme case, there should be no updating. If all firms were alike in their propensities to explosions, then having had an explosion would say nothing about risk in the future. This might be true, even though ordinary losses of lives might be well explained by an updating distribution, such as the beta binomial.¹² More realistically, the explosion probability could be fairly well known, in the sense that experience would not much change one's predictions.¹³

But if catastrophes should count less in experience, I suspect that it is not because the probability of the catastrophe should not be adjusted. Rather, it is the size of loss that does not tell us much. When New York's World Trade Center was bombed using a mixture of fertilizer and petrochemicals, the number of fatalities fortunately was limited. But the message was clear: even relatively unsophisticated groups can and have made massive bombs from readily obtained materials. A bomb of the same type in Oklahoma City led to much more catastrophic consequences. It shook America, causing a welfare loss far beyond the lost lives and mangled bodies. (The property loss for Oklahoma City was comparatively minor.)

Such stories suggest that catastrophes should often be thought of in terms of compound probabilistic models. There is a probability of loss and a size of loss.¹⁴ It is quite possible that the size of loss is a chance event for which we update little, but events offer substantial learning about the probability of loss. If so, then in pricing for the future, we should count strongly that a catastrophe occurred, but the precise number of people killed may not tell us much. This is not the place to rehearse such models mathematically. I will just note that the loss of lives might be exponential with the key parameter known or virtually known, but with the probability p subject to a beta binomial with relatively small values for m and n , implying substantial learning. The critical question may be, when you learn that an insured had a catastrophe and many people were killed, should you learn more about the risk of catastrophe or more about the number of people killed given a catastrophe?

Needless to say, this probabilities-versus-quantities question is relevant in many circumstances where loss is not measured in human lives. When there is an oil spill at sea, what

frequently determines whether it is minor or disastrous is proximity to shore, the weather, and the intensity of the waves. If the insurance premium already accounts for the size and route of the tanker and whether it is double-hulled, relatively little may be learned from the amount of damage. Somewhat more may be learned about the particular company's risk of loss, perhaps because its captains are poorly trained.¹⁵

Catastrophes are sufficiently low-probability events that for all but the most common catastrophes (say multiple deaths in an auto crash) it would probably be difficult to estimate which model best predicts the likelihood of losses of various magnitudes. This suggests that insurers might undertake a form of meta-analysis, looking across catastrophes. From this we might learn, for example, to what extent we should update our probability assessments from an occurrence and how we should update our estimates on size of loss. The updating would come in two parts. First, we would like to specify the likelihood that any particular insured will suffer a catastrophic loss. Second, we would secure information about the risk of catastrophic loss for a class of insureds.

Some catastrophes are so grave that once one is suffered, the insuring entity is put out of business; future insurance rates are not a concern. This could happen, for example, with a physician who is found to be grossly negligent or a company whose major product proves carcinogenic. However, insurers must decide what to charge firms such as Exxon after the Valdez spill, USAir after its era of many crashes, a Kobe business establishment that suffered massive losses in the earthquake, or the World Trade Center after being sabotaged. In adjusting rates for such experiences, it may be excessive to fully count the magnitude of catastrophic losses. If so, the Equitable was right to underweight catastrophes in setting premiums and dividends.

The process of updating risk estimates after catastrophes requires a refined mix of judgment and actuarial skills. Most of our methods and our thinking are designed for a world that is relatively stable. Hence, we have a great deal of experience with one bad outcome out of a long string. Judgment plays a predominant role in recognizing when the catastrophic event signals that the world has changed. For example, after a couple of plane crashes due to metal fatigue it is important to determine whether the aging of the fleet has altered risk factors for the worse. When the first big malpractice awards come in against accounting firms, that probably signals a sea change. The World Trade Center and Oklahoma federal building bombings probably both represent and help instigate a sea change in major terrorist activities in the United States.¹⁶ However, the Taegu, Korea, gas blast, which took roughly as many lives in the week following Oklahoma, should not noticeably increase risk estimates of catastrophes. It provided little new information, was an extreme outcome given the nature of the triggering event, and if anything will promote safer activities. Some catastrophes may on net signal lower risks in the future as we learn how to deal with them and take expensive actions to reduce their likelihood and consequence—in part because the catastrophe strengthens political will to make such expenditures. For example, in response to significant damage in a recent earthquake, Stanford University undertook expensive measures to reinforce its earthquake-prone buildings.¹⁷

4. Mitigating catastrophic losses

The concept of moral hazard is well understood: insured individuals will take less care—in effect spend fewer personal resources—to reduce the likelihood of a bad outcome. Little

discussed, at least in economics, is another distortion due to insurance. Insurance can dramatically affect the insured's choice of distribution of outcomes. We shall refer to these two types of moral hazard as effort discouragement hazard and distribution distortion hazard.¹⁸ Both of these effects play a major role in the way insurance should be structured to deal with catastrophes.

4.1. Distribution distortion hazard

The distribution distortion (DD) hazard occurs when an individual at risk can take actions that affect the tradeoff between the probability of a bad outcome and its consequences. A manufacturing company receiving slender evidence that its product may be hazardous faces such a choice: if it pulls the product, it incurs a sure but contained loss. If it does nothing, it risks a larger catastrophe in the unlikely circumstance that the hazard is real.

Actions that trade off probability and size of loss are common in financial contexts. For example, a developing nation has an unpromising balance sheet. The decision maker (DM), an international creditor, can refuse to roll over a loan and expect a loss. Alternatively, it can provide new funding, relying on the finance minister's assertion that matters will improve. If they don't, the loss will be far bigger.

Ideally, an individual pursuing a positive expected value activity—that is, one with bright future prospects—would plow ahead. But we know from events, ranging from the Barings debacle in Singapore to the widespread collapse of S&Ls in the United States, that it is frequently worthwhile for the DM to go forward even when the expectation is negative. (Thus, after Mr. Leeson had lost a great deal of money, he gambled some more, hoping that the Japanese market might move strongly in the right direction.) In both of these instances, outside parties were insuring part of the loss—respectively, the Singapore Exchange, and agencies of the U.S. government.¹⁹ Agency problems also fostered poor outcomes in these instances. Mr. Leeson, or the president of a savings and loan institution, can be fired only once; for him losing \$100 million may be almost as bad as losing \$300 million.

Basically, in both cases, DMs exchanged a high-probability low-consequence event for a lower-probability but high-consequence event. (After all, the recovery efforts might have been successful.) In both cases, I suspect that an *ex ante* analysis would reveal that the continuation of efforts could be expected to increase the mean of the loss. They certainly increased the variance.

Physical catastrophes may also be subject to actions that increase both the mean and variance of losses. The recently refurbished QE2 set sail this spring with repairs incomplete and a hope that all would work out. But severe plumbing problems, including leaks and exploding toilets, turned its maiden voyage into a nightmare. Cunard lost a small fortune in compensation payments to passengers and a large one in loss of good will. The option of cancelling the trip, though offering a certainty of loss, would have incurred a much smaller expected and actual loss.

Let me illustrate with a hypothetical company that owns a contaminated waste site. It can either sit with it, hoping that the toxic materials will not seep into the groundwater, or it can clean it up. Sitting, rather than cleaning, offers a smaller probability but a higher consequence of catastrophic loss.

Table 1. Options for a company owning a contaminated waste site.

| | Probability of Incurring Cost | Discounted Value of Cost | Discounted Expected Cost |
|----------|----------------------------------|-----------------------------|-----------------------------|
| Clean up | 1 | 100 | 100 |
| Sit | 0.2 | 1000 | 200 |

Say the parameters of the case were as shown in Table 1. Even if all parties were risk neutral, cleaning up the mess now would be an actuarially favorable activity. But let us say instead the company was insured. The policy might have a 200 deductible with a 0.1 copayment above that. Now if the company sits and the waste seeps, the discounted cost to the company is 280. The discounted expected cost of sitting would be 56, so the company might well take that strategy. (It is unlikely the insurance would pay for the cost of a clean-up before any leakage occurred.)

In this example, one action could lead to one outcome, the other could lead to two outcomes. In many situations, there will be a distribution of outcomes for any action. The alternative choices of the DM each will offer a different distribution of losses. The presence of insurance could lead to a substantial distortion in the choice of distribution; hence the label distribution distortion hazard. The DM makes a distorted choice of distribution—increasing both the mean and variance of loss—due to the presence of insurance.²⁰

4.2. How can insurers confront distribution distortion?

Assume, as is traditional in insurance problems in economics, that the DM is risk averse and the insurer is risk neutral. Assume further that the DM cannot commit to take a particular action and that if his action is monitorable it is not punishable. The insurance company charges a fair actuarial rate, taking into account the actions the insured DM will take in light of the insurance schedule. In effect, the insurer and the insured begin with a common interest in selecting the optimal schedule. Once that schedule is agreed upon, however, the insured will have an incentive to exploit the insurance relation by engaging in distribution distortion.²¹

This is an agency problem, of course: the insured does not take into account the welfare of the insurer. Quite apart from insurance, agency problems already present make catastrophic outcomes more likely. For example, the manager of the company can be fired only once. His penalty may not be much more severe if the company loses \$200 million due to his folly rather than a mere \$100 million. So too, bankruptcy laws, which truncate the bottom of the distribution, make catastrophes more likely. If cigarette companies are ever made to pay for the health risks of cigarettes, the companies' assets but not their investors' assets are at stake. It is no surprise that when companies do go bankrupt, their net asset position is likely to be substantially, not slightly, below zero. This implies that they had an incentive to gamble at the end and took it.

If distribution distortion were the only problem, it could easily be dealt with through the *sliver solution*. This approach merely charges the insured a tiny, constant fraction of

the losses—say, 0.1 percent. This essentially transfers all significant risk to the insurance company but fully aligns the interests of the insured and the insurance company. Distribution distortion will not be a concern.

There are at least two reasons why the sliver solution is not employed: first, some costs of dealing with the risk—such as the cost of better monitoring employees at Barings—would not normally be insurable and compensable. Second, effort discouragement hazard will rear its head if the insured bears but a small portion of the loss. Therefore, to deal with this hazard, the insured must have a significant amount of expected utility at stake due to uninsured risk costs. In essence, moral hazard creeps in when insurance coverage is too complete.²²

Our goal is to understand the nature of the optimal insurance when both distribution distortion and effort discouragement hazard are present. To get our thinking straight, let us start with merely effort discouragement hazard. The insurance arrangements basically yield two outputs—costs to the insurance company and utility loss to the individual.

4.3. *Effort discouragement hazard*

Let us assume that effort discouragement affects only the probability of loss. Hence, we have $p = P(e)$, where p is the probability of an accident and the distribution of the cost of the accident is independent of e . The problem, in essence, is to find the optimal expected utility loss, call it U^- , associated with uninsured risks, while having the insured absorb as much expected dollar losses as possible. The solution to this constrained-optimization problem is to employ a deductible, D , the amount the insured will lose whenever there is a loss greater than or equal to D . When there is no loss, or a loss less than D , he will merely bear all the cost.

We are comfortable with the concept of deductibles. Among their other virtues, they reduce administrative costs. Moreover, as Arrow showed, if a certain sum must be paid to the insurance company—for profits or overhead—the cheapest way to raise these funds, in terms of lost utility to the insured, is through a deductible. It extracts money from the insured when the utility cost is cheap.

But if the insured has some ability to shift the shape of the distribution, a deductible may be a far from optimal approach. If the DM can reduce the probability of a loss greater than D , he saves D in direct proportion. He will merely choose the distribution giving the least likelihood of a loss greater than D . That is, he will trade bigger losses for a smaller probability. He will seek to turn distributions with unfavorable outcomes into those that include catastrophes.²³

We shall deal with two separate cases, both of which show that conventional wisdom about the form of ideal insurance may be misguided. In the first case, an insured's effort reduces the probability of an accident. In the second, it shrinks any loss proportionately. We assume constant or decreasing risk aversion, as is normal.

4.4. *Optimal insurance given insured's potential actions*

The form of an optimal insurance arrangement will depend on many factors. Here we focus on the insured's ability to influence losses. We assume the insurer is risk neutral and the

Table 2. The form of optimal insurance policies given insured's potential actions.

| Role of Effort | Distribution Distortion | |
|-----------------------|---|---|
| | No | Yes |
| No role | A. Complete insurance | B. Sliver solution |
| Affects probabilities | C. Deductible | D. Insured's share of uncertain loss may be greater for either high or low losses |
| Affects size of loss | E. Insured's share of uncertain loss may be greater for either high or low losses | F. Insured's share of uncertain loss may be greater for either high or low losses |

insured is risk averse. Table 2 summarizes the nature of the optimal plan for six sets of conditions.

Absent both distribution distortion and a role for effort (*A*), the optimal policy is complete insurance. If effort plays no role, but there is distribution distortion (*B*), the optimum plan merely charges the insured a tiny fraction of whatever loss is incurred, what we label a sliver solution. If all potential losses are finite, the sliver can be made small enough to make the insured risk neutral about losses (since his utility is effectively linear over the relevant range). Then the insured's incentives will align perfectly with those of the risk neutral insurer, and he will select the efficient distribution. If effort affects the probability of loss (*C*), a deductible is the preferred mechanism, since it most evenly spreads losses for any potential level of incentive, hence sacrifices the least value to risk aversion.²⁴ These are the easy cases.

If a choice under the control of the insured affects the distribution of losses, conditional upon there being a loss, matters become more complicated. Two numerical examples are presented in the appendix to address cases where the insured can affect the distribution of losses. In the first example, effort affects the probability of an accident, and there is distribution distortion (*D*). In the second, effort shrinks any loss proportionally (*E*), which though not distribution distortion gives the insured some control over the loss distribution.

These examples, though plausible in structure, lead to unfamiliar results. Consider two alternative insurance policies, which in some sense involve opposite approaches. The first employs a deductible, a standard insurance arrangement. The second places a cap on insurance company payments, meaning the insured bears all losses beyond some point. For both numerical examples, the best cap policy is preferred to the best deductible policy, which is surprising.

The optimal unconstrained scheme for *D* has some features of both a deductible and some of a cap: it makes the insured responsible for initial losses up to a limit but then makes him responsible for a larger amount when losses reach a level that would be likely given an undesirable choice of the distribution of losses. Among five plans considered for the second example, *E*, where effort shrinks the size of loss, the best arrangement involves a cap on insurance expenditures.

Our two examples show that the optimal insurance plan can require the insured to pay a larger share of high than low losses for cases *D* and *E*. Different numerical values for the same cases can lead to the opposite result; it would be optimal for him to pay a larger

share for low losses, as happens with ordinary insurance. That outcome results when both effort and distribution distortion are minor considerations. For cases *D* and *E*, therefore, we find that the insured's optimal share of the loss can increase or decrease with the size of loss. Since *E* is a special case of *F*, this result applies for that case as well. Most traditional analyses of insurance address only boxes *A* and *C*, ignoring the important possibilities of distribution distortion, and that effort can affect the size of loss.

Conventional wisdom suggests that insurance company payouts should be concentrated to protect against catastrophes, implying payouts will be highest proportionally when losses are great. Because of this, we tend to see deductibles as a major feature of insurance policies that both maintains the insured's effort and reduces administrative costs. (Policies may also incorporate some sharing of losses by the insured to maintain incentives, say if effort can control losses once they occur.) Yet if effort has its greatest effect on losses at the high end or if distribution distortion is a threat, then, as the appendix shows, the insured's optimal share of losses might well increase with the magnitude of loss.

4.5. *Mitigation in action*

If the insured can take actions that shift the distribution of losses, insurers that fail to impose significant penalties on insureds for big losses are sure to see such losses occurring more often. Distribution distortion is a significant problem, particularly when human action—as, for example, with embezzlement, staving off bankruptcy, or dealing with unsafe products—is the critical factor affecting payoffs. Given distribution distortion and the possibility that effort affects the size of losses, insurance plans with caps on payments to the insured, or other severe penalties for big losses, merit serious consideration.

The goal of emplacing incentives to mitigate catastrophes is more than a theoretical nicety. In the United States, policy reform on disaster relief—which cost \$120 billion (1993 dollars) from 1977–1993—is expected to place increased emphasis on hazard mitigation as well as increased use of insurance as a means to reduce the costs of federal disaster assistance. The push for mitigation, as a recent Senate Task Force observes, starts with the observation that “many hazard losses are preventable.” Beyond direct regulation, its report discusses strengthening incentives for mitigation by pricing for risk, or subsidizing mitigation expenditures, but does not consider structuring insurance policies to maintain incentives (U.S. Senate Task Force, 1995, pp. xiii, 60–74).

Whether the government itself can be an effective insurer is open to question. Its insurance policies tend to be subsidized with little or no pricing for risk. Such programs as flood and crop insurance have not attracted widespread participation. The alternative proposal of private insurance with government reinsurance could result in an “open-ended obligation of the federal government and a federal guarantee of the solvency of the property/casualty insurance industry” (U.S. Senate Task Force, 1995, p. 74).

It is widely recognized that current insurance and disaster assistance policies create insufficient incentives to limit losses in catastrophes. The challenge to the insurance industry, government, and insurance scholars is to determine the mix of policies that can effectively strengthen these incentives, while still providing substantial risk protection to insureds.

5. Conclusion

Catastrophes play a central role in the simple theory of insurance. In that theory, catastrophes are the principal insurable event, representing uncontrollable low-probability high-consequence outcomes. I have argued that catastrophes are not mere bolts from the blue against which the insured has little protection; they are much richer phenomena.

I have identified three aspects of catastrophes, each of which suggests that our treatment of them should be modified. The first aspect is that catastrophes may often be solely or predominantly the consequence of human action. In such circumstances, the payment to a claimant might appropriately be higher the greater the number of claimants who incurred similar losses. The second aspect is that the mean value of losses may be a poor indicator of the future consequences of catastrophes.

The third aspect is the ability of insureds to affect the distribution of losses. Insureds can frequently reduce the probability of loss if they are willing to incur larger losses; we call this distribution distortion. Under such circumstances, insurance policies may have to be structured reverse to the way we ordinarily see, with greater risk sharing at the top not bottom.

Catastrophes appear more frequently when potential victims have insurance than when they don't. They are complex and can easily be influenced by the form of an insurance policy. Understanding the anatomy of catastrophes will help insurers write policies that strike a sound balance between appropriate incentives and protection against risk.

Appendix: The form of optimal insurance plans²⁵

This appendix demonstrates some assertions in the text about the nature of the optimal plan when the insured makes an effort decision. We deal with specific parameter values. Utility takes the separable and additive form

$$U(w, E) = -\exp(-w) - aE,$$

where E represents the effort level, a is a parameter, and aE is the cost of making an effort. In our first example, effort affects the probability of loss; in the second, it affects the magnitude of loss.

Distribution distortion hazard

The insured can select between two different distributions of loss, A and B . With A , the magnitude of losses is uniform between 0 and 2. Distribution B cuts off the losses but produces a higher average loss when a loss happens. If a loss happens with B , one-fourth of the time it will be 0; the other three-quarters of the time it will be uniform between 1 and 2. The expected losses for A and B are 1 and $9/8$ respectively. Obviously, A should be chosen; it offers smaller expected losses. Moreover, B suffers from a greater variance of payoffs.

We simplify by assuming that effort is discrete, and either high ($E = 1$) or low ($E = 0$). With low effort, the probability of loss is .4; with high effort, it is .1. The sequence of moves is as follows. The insurance company offers a fair actuarial plan, taking account of the insured's expected choice of effort level and selection of A or B . However, these choices are not monitorable and cannot affect the premium the insured pays. After the insured pays the premium and chooses his effort and distribution, the lottery on losses is run and any acquired insurance payments are made. The insured cannot be required to bear more than full losses; otherwise he would not file a claim.

Post-lottery wealth will be

$$w = w_0 - x + I(x) - P(I),$$

where w_0 is initial wealth, x is the magnitude of loss (if any), $I(x)$ is the insurance payment for loss x , and $P(I)$ is the equilibrium premium given $I(x)$. Our parameter values are $w_0 = 3$, and $a = 0.00428$.

We consider six insurance plans. Their order from worst to best from the standpoint of the insured is no insurance, deductible, cap (insured pays all losses beyond a fixed point), proportional (linear sharing of costs by insured and company), rectangle (insured pays fixed amount for any loss above specified level), and slant-step (see Figure 1). (The insurance company always breaks even and hence is indifferent among the plans.)

It is easily verified that high effort is efficient. The parameters for each plan were selected to just induce the insured to put forth high effort. Figure 1 shows the optimal form of each plan, with parameter values and certainty equivalent (CE) losses. The CE loss is the amount the individual would be willing to lose for sure, assuming low effort, as opposed to being subjected to the vagaries of the real world. For the deductible plan, the insured selects distribution B ; for all other plans he chooses A . The design of the insurance plan has three goals: (1) to spread the insured's uninsured losses, (2) to encourage an efficient level of effort, and (3) to concentrate the insured's losses where total losses are great so as to curb distribution distortion hazard (prevent the choice of B).

The major lesson of this example is that optimal insurance plans may look unfamiliar and that when distribution distortion is a concern, incentives against large losses must be maintained.²⁶

Effort affects the magnitude of loss

Similar reasoning may apply if effort differentially reduces large losses. Let us now assume that the probability of a loss is .1, independent of effort, but that effort reduces the proportionate size of the loss. If there is a loss, x represents what the loss would be without effort, where x is uniform on $[0, 1]$. For this example, effort is permitted to be continuous. The loss after effort is $x(1 - E^{.5})$. For loss x , wealth will thus be

$$w = w_0 - x(1 - E^{.5}) + I[x(1 - E^{.5})] - P(I).$$

The utility function is as before, but $a = 0.05$. Initial wealth is 3.

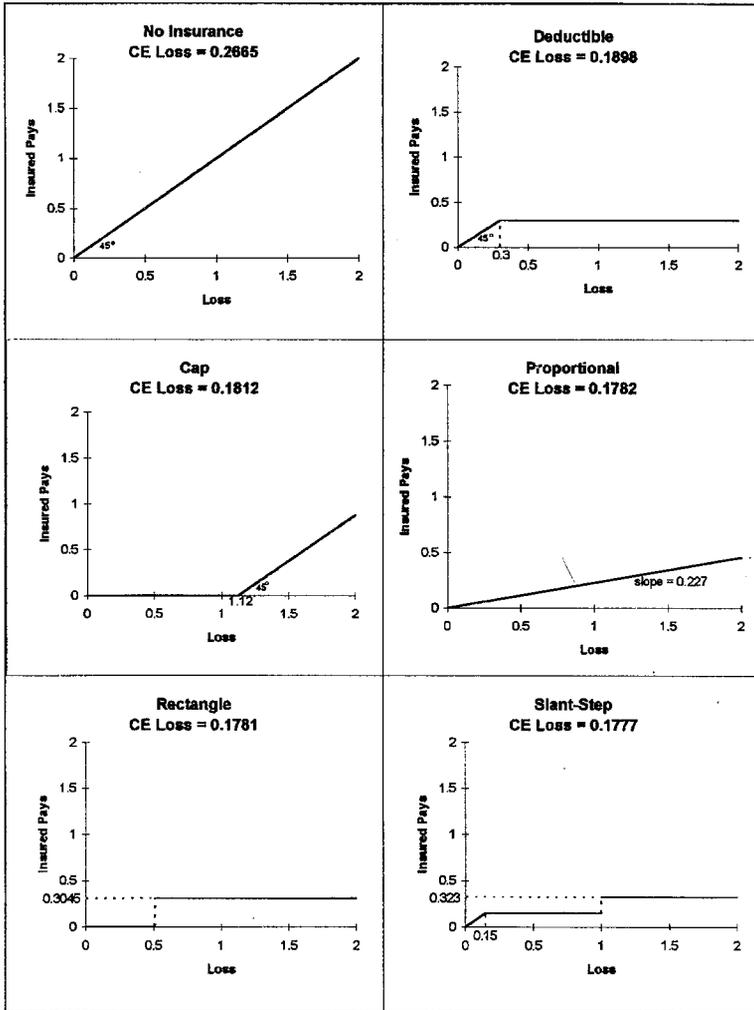


Figure 1. Optimal insurance plans with effort and distribution distortion.

For this example, the rationale for putting some burden on the insured for losses at the high end of the scale is that this is where effort has the greatest effect on losses. We consider the same plans as before, excluding the slant-step (a plan optimized for the first example). Table 3 reports their performance. (Effort and certainty equivalent losses are multiplied by factors for convenience in presentation.) The ordering of plans from worst to best is shown in Table 3.

Table 3. Ordering of plans.

| Plan | Parameter Values | Effort $\times 10^5$ | CE Loss $\times 100$ |
|-------------------------|--------------------------|----------------------|----------------------|
| No insurance | | 256.87 | 6.7130 |
| Deductible ^a | $d = 0.00+$ | 0.00+ | 5.0000 |
| Proportional | $q = 0.04$ | 0.11 | 4.9971 |
| Rectangle | $t_1 = 0.60, t_2 = 0.03$ | 1.00 | 4.9868 |
| Cap | $c = 0.83$ | 9.02 | 4.9683 |

a. Initial effort has an extreme benefit/cost ratio given its infinite derivative at zero. Hence effort will always be put forward when the deductible is not zero. However, a deductible is a highly inefficient structure when effort reduces losses proportionally. The optimum deductible is tiny, below .0001. At the optimal deductible, effort is several orders of magnitude less than with any other case. The CE loss $\times 100$ falls below 5.0 in the fifth decimal place.

Among these simple plans, the best one places a cap on what the insurance company pays out.²⁷ The insured pays all losses above 0.83. This focuses his monetary incentive where his effort has the greatest effect. His effort proves significant relative to the other insurance plans.²⁸

Summary

If the insured can tilt losses toward catastrophe, or if his efforts will do the most to reduce catastrophes, either by reducing probabilities or reducing magnitudes, then traditional lessons about the form of optimal insurance are likely to be wrong.²⁹ Even if the insurance company is risk neutral and has infinite resources, it would keep the insured at some risk when losses are high. And the insured, who in our formulation reaps any savings in the form of reduced premiums from stronger incentives for care, has the same interest as the insurer in finding the optimal form of policy.

Notes

1. Of course, a few more Kobes might be a serious problem to an insurance company; such a situation would represent a notable change in well-estimated probabilities of natural events. There is much talk at present that global warming may have increased the likelihood of weather-related catastrophes, such as hurricanes and droughts.
2. The duration of unemployment insurance (UI) in the United States is often adjusted for economic conditions. With unemployment insurance, moral hazard arises in a second form: will individuals conscientiously seek jobs? There is strong evidence that the ending of UI is associated with a return to work [Meyer, 1990]. Here too there is justification for continuation of payment when unemployment rates are high; it is less likely that the unemployed would be able to find jobs. Some economists propose a macroeconomic justification for generous UI in a recession—to maintain aggregate demand.
3. Daniel Kahneman, the distinguished decision theorist who lost his house in the disastrous 1991 Oakland, California fire, reports that insurance companies were extremely generous in their compensation. They were

lax in requiring documentation of losses and in some cases paid claimants more than the total value of their insurance. Kahneman believes that great media attention and relatively easy comparisons among insurance companies promoted this generosity.

4. It would be difficult to have a plan that paid off less for isolated losses. Ironically, however, insurers might be able to formulate a plan that eliminates some cost sharing by insureds when there was a common catastrophe. This is not unlike the observation that supermarkets could not implement a premium check-out line where one paid more to check out quickly but could put in place a reduced price line for those willing to endure a slower checkout.
5. In the United States in early 1995, when the Republican congressional majority first came to power, it was expected to legislate substantial limitations on liability penalties. At time of writing, in late 1995, this outcome seems much less likely. Appropriate insurance writing decisions had to gauge the likelihood of significant liability reforms becoming law. Republican congressional majority will legislate reduced liability penalties.
6. Nutter addresses two of the most troubling areas in insurance in the United States, natural disasters and environmental liability under Superfund. He identifies efforts to address these issues through joint undertakings of insurers and government. He observes that if these efforts are not successful, the economies "will force public solutions" [Nutter, 1994].
7. Financial markets can sometimes fill the void. Derivatives written on stock indices have been great successes. The market makes the requisite actuarial calculations. Surprisingly, a contract on the Consumer Price Index in the United States was a failure.
8. With government regulation, of course, private insurers are often prevented from taking certain pricing approaches based on sound actuarial principles. Auto insurance in some states in the United States has led to catastrophic losses for insurers because both rates and rating systems have been severely constrained, and also because regulators have imposed levies on insurers to cover deficits in assigned risk pools.
9. In April, the U.S. Commissioner of the Currency gave a major speech in which he warned against risk-increasing actions of banks, which were making riskier loans. This somewhat technical speech was reported as a lead article in the *New York Times* [Bradsher, 1995, p. 1].
10. Indeed, even if other firms were fully informed, they might take a lesser profit on the bad-experience firm, since it would not be establishing a premium precedent for other insureds.
11. Lemaire, [1995, pp. 71–85] computes the elasticity of thirty different national bonus-malus systems, where elasticity measures the long run response of the premium to the change in claim frequency. Ideally, the elasticity should be 1. His highest elasticity was .45 for Switzerland; for two-thirds of countries elasticities were below .20.
12. With the beta binomial, the probability that an individual dies in the next year is given by m/n . These numbers are updated on the basis of experience. For example, if we started with $m = 50$ and $n = 1,000$ in a 100-person firm, and if nine people die the next year, the updated probability would be 59/1,100.
13. If for the beta binomial it was equivalent to 50/5,000, then having an explosion the following year would be 51/5,001, a smidgen less than a 2 percent increase.
14. Lemaire [1995, p. 205] observes that "all bonus-malus systems in force throughout the world [except Korea] penalize the number of reported claims, without taking the costs of such claims into account. A mere scratch causes the same premium increase as a serious bodily injury accident." His analysis shows that "the malus applied to a single bodily injury [expensive] claim should be as high as from four property damage claims" [Lemaire, 1995, p. 210]. Lemaire's result derives because high claims predict high claims, not because they indicate a higher probability of a loss.
15. Sometimes we can choose the model. When looking at commercial airline risk, frequency of fatal accidents per mile of flight, with takeoffs and landings counting substantially extra, is more significant than numbers of lives lost, which will depend significantly on whether a doomed plane was full or empty, since load factor has little effect on risk. *Newsweek* provided a cover story on airline safety in its April 24, 1995, issue. It utilized experts' estimates of comparative risk on a per-flight basis. Length of flight was not included as a risk factor, presumably because most accidents occur on takeoffs and landings.
16. The Oklahoma tragedy should update risk assessments of losses to terrorism. How those updates should be broken down between learning that some existing groups are dangerous, that massive bombs can be constructed from everyday materials, and that such bombs can kill in the hundreds is not clear.
17. Professor Kenneth Scott, private communication [1995]. Since the Oklahoma City bombing, the government has undertaken a quite different effort to reduce catastrophic losses. It has initiated vigorous measures, including some curtailing civil liberties, to reduce terrorist attacks.

18. Habits and norms are often useful in controlling moral hazard behavior. Insurance, unfortunately, tends to encourage "bad behavior" and in time erodes the norms that were initially helpful in constraining such behavior. This is true whether the insurance is privately or publicly provided. On the erosion dangers from public insurance provision, see Lindbeck [1995].
19. Frequently the insurance arrangements are only implicit. For example, banks lending to Mexico in recent years perhaps suspected that in a peso crisis the U.S. government would provide some sort of bailout, as it did. Sometimes the "insurance" comes from the forbearance of regulators, which do not make financial institutions take writeoffs. U.S. banks were accorded this privilege in the Third World debt crisis of the early 1980s. Japanese banks receive that treatment today, with their underwater real estate loans.
The potential for bankruptcy, which limits downside losses for shareholders, offers companies a form of insurance against catastrophes. When confronted with massive toxic tort suits, Johns-Manville declared bankruptcy to protect against asbestos liability; A.H. Robins did the same for the Dalkon Shield, and Dow Corning is now pondering filing for bankruptcy as part of its strategy for dealing with lawsuits over silicone breast implants [*New York Times*, 1995, pp. D1-D5].
20. A linear slice insurance plan—say, one bearing 50 percent of all costs—may induce moral hazard, but it would never induce the decision maker to pick a higher expected cost outcome. It could lead to a lower expected costs outcome, if risk aversion had previously led to expensive hedging.
21. He would like to commit not to engage such a nefarious activity but cannot do so if his actions are not monitorable and hence can be neither constrained nor priced.
22. For a wide range of catastrophic-loss situations there are categories of loss that should not be insured, in the sense that even if available at actuarially fair rates, the insured should not seek to make himself whole. For example, a lawyer who is a ski buff should not insure his leg, despite the enormous loss of utility should he no longer be able to ski. Efficient insurance keeps marginal utility constant across states, not absolute utility. With efficient insurance, therefore, there will often be collateral uninsured losses, which tend to reduce moral hazard. Moral hazard is more of a problem for warehouses, which burn at night, than for light planes, which take pilots with them when they crash.
23. Interestingly, given risk aversion, the second-best solution, as would be achieved if his choice of distribution but not his effort were both monitorable and punishable, would be to accept a greater probability of loss in return for a smaller D . Starting from the mean-minimizing distribution, this is a move in the opposite direction from what DD induces with traditional insurance.
24. Zeckhauser [1970] discusses additional traditional insurance arrangements. These are the easy cases.
25. Todd Olmstead expertly conducted the calculations for this appendix using Microsoft Excel.
26. For this example, the ratio of densities of high effort and low effort losses is constant for losses above 1. Hence, the optimal plan does not increase the insured's payment for losses above 1.
27. If effort were monitorable, pricing could achieve a first-best solution. The optimal effort would be 68.31×10^5 , the CE loss would be 4.9358×100 , and there would be full insurance.
28. When there is no insurance, not surprisingly, he works much harder. The reader may prefer to think of a rescaled E as effort squared or cubed, with an accordingly increased marginal cost of effort. Given that interpretation, the disparities between work efforts across plans diminish dramatically.
29. Here I merely provide numerical examples. I leave it to mathematical wizards to derive general conditions on the optimal form of policy given various assumptions about the insured's opportunity to make choices about effort and the distribution of losses.

References

- BRADSHER, K. [1995]: "Banking Regulators Taking Close Look at Lending Risks," *New York Times*, April 9, p. 1.
- LEMAIRE, J. [1995]: *Bonus-Malus Systems in Automobile Insurance*. Kluwer, Boston.
- LINDBECK, A. [1995]: "Welfare State Disincentives and Endogenous Habits and Norms," mimeo, Institute for International Studies, Stockholm, April.
- MEYER, B.D. [1990]: "Unemployment Insurance and Unemployment Spells," *Econometrica* 58(4) (July), 757-782.
- NEW YORK TIMES. [1995]: May 6, pp. D1-D5.

NEWSWEEK. [1995]: April 24.

NUTTER, FW. [1994]: "The Role of Government in the United States in Addressing Natural Catastrophes and Environmental Exposures," *Geneva Papers on Risk and Insurance: Issues and Practice*, 19(72) (July), 244-256.

U.S. NATIONAL SAFETY COUNCIL. [1994]: *Accident Facts*. U.S. Government Printing Office, Washington, DC.

U.S. SENATE TASK FORCE ON FUNDING DISASTER RELIEF. [1995]: *Federal Disaster Relief*, U.S. Government Printing Office, Washington, DC.

VISCUSI, W.K. [1995]: *Reforming Products Liability*. Harvard University Press, Cambridge, Mass.

ZECKHAUSER, R. [1970]: "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, 2(1) (March), 10-26.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.