

## Chapter VI

### Risk

Life, by its very nature, is risky. Most people do not like risk, yet reducing risk is expensive. Individuals, businesses, and governments must make tradeoffs between the costs of living in a risky world and the costs of reducing that risk. Everyone through their behavior indicates that they are willing to take on certain risks in their lives. The fact that people ride in automobiles, as opposed to riding in tanks or armored vehicles which are much safer, indicates that individuals make tradeoffs between the cost and comfort of an automobile and a safer mode of travel. In general, individuals engage in risk avoidance activities up to the point where the marginal benefit of the reduced risk is equal to the marginal cost of reducing the risk. The fact that people do not spend the resources to make their life risk free suggests that life itself is not priceless.

This chapter covers several ways individuals and society attempt to deal with risk. After a review of the economics of uncertainty and risk in Section A, the next two sections consider essentially private, market-based ways of dealing with risk. Section B covers the demand for and supply of insurance. Section C is an overview of the manner in which market prices for products, services, and jobs adjust to reflect the associated risk. Section D analyzes how tort law deals with accidents and injuries that result from a non-contractual, non-negotiated relationship. Finally, Section E presents an overview and critique of federal risk regulation.

#### A. Economics of Uncertainty and Risk

Individuals and firms must make decisions when they are not certain of the outcome. However, they often have some idea of the range of possible outcomes. This section explains how the size of the range of possible outcomes has an impact on decision making. The first subsection introduces basic concepts from probability theory which are then used to calculate some descriptive statistics — expected value and variance — regarding the range of possible outcomes for a decision. The next subsection describes the assumption that individuals are risk averse and explores some implications of that assumption.

##### 1. Basic Probability Theory

Some theoretical economic models — for example, the perfect competition model in Chapter VII — rely on the assumption that individuals have perfect information regarding the outcome of any particular decision. Such decisions have **certain outcomes**. A more realistic assumption is that individuals contemplate a variety of potential outcomes and assign relative likelihoods to their actual occurrence. In such cases, the decisions have **uncertain outcomes**. When decision makers have information regarding possible outcomes and their relative likelihoods, the tools of probability theory can be used to analyze economic decisions.

A **probability** is a number between 0 and 1, inclusive, that expresses the likelihood that some specific event will occur. Some probabilities can be estimated through a process of experimentation. For example, we would expect that the probability of getting "tails" as a result of flipping a fair coin is 0.5 (which is equal to 1/2 or 50%). When probabilities can be established in this fashion, they are **objective probabilities**. In other cases, probabilities are estimated on the basis of a subjective "best guess" or one's prior experiences. When a particular individual's opinion enters the estimation, she is using a **subjective probability**.

It will often be useful to identify situations in which two probabilistic events are

independent of each other. Events are **independent** when the probabilistic outcome of one decision does not affect the probabilistic outcome of a second decision. For example, while driving through Missouri on a cross-country road trip, Jerry decides to purchase a state lottery ticket. This decision over the use of funds has an uncertain outcome to which probabilities can be assigned. After purchasing the Missouri state lottery ticket, Jerry continues his trip across the country. While in Kansas, Jerry decides to purchase a lottery ticket for that state's lottery. However, by purchasing a lottery ticket in Kansas, Jerry has not changed the probability that he will win the lottery in Missouri. Thus, the probabilities for winning either lottery are independent of each other. Note that this is different than estimating the probability that Jerry would win both lotteries. The probability of two independent events both occurring is simply the product of the probabilities. Thus, if Jerry had a 25% chance of winning the Missouri lottery and a 25% chance of winning the Kansas lottery, then his chance of winning both lotteries would be  $25\% \times 25\% = 6.25\%$

For purposes of analysis, assume that outcomes and probabilities are both **complete** and **mutually exclusive**. Completeness suggests that while the decision maker does not know the specific outcome of any decision, she is aware of all potential outcomes. Mutual exclusivity indicates that the result of a particular decision will be x or y, not x and y. These two assumptions allow the use of the notion of an expected value in analyzing uncertain decisions.

## 2. Expected Value and Variance

Given a set of objective or subjective probabilities and all possible outcomes, the expected outcome of an uncertain decision can be calculated. The **expected value** of an uncertain decision is the product of outcomes and their probabilities summed across all possible outcomes:

$$E(x) = P_1(x_1) + P_2(x_2) + \dots + P_n(x_n)$$

where  $E(x)$  is the expected value of  $x$ ,  $P_i$  is the probability of any outcome  $I$ , and  $x_i$  is the actual value of outcome  $I$ . Each  $P_i$  must be between 0 and 1, inclusive; and the sum of all  $P_i$ s must equal 1. This sum provides the likely value of potential outcomes, where only one of the outcomes will occur.

Consider the following example. Suppose that you are representing a widower in a wrongful death suit. The case has just gone to the jury, and the defendants offer to settle for \$1,000,000. In deciding whether to accept this offer, you must consider the amount of damages you expect to collect by waiting to hear from the jury. Based upon your experience as an attorney and the willingness of the defendant to settle at this point in the litigation, you feel quite confident that the plaintiff will be awarded a substantial judgment. Moreover, because the trial is complete, your costs in terms of time and resources are the same whether you settle or wait for the jury. You feel that there is a 30% chance that the jury will award \$750,000, a 50% chance that the jury will award \$1,000,000, and a 20% chance that the jury will award \$2,000,000. Thus, the expected value of the jury award is:

$$E(x) = 0.3(\$750,000) + 0.5(\$1,000,000) + 0.2(\$2,000,000) = \$1,125,000$$

By your calculation, the expected result of the jury award exceeds the settlement offer by \$125,000. In other words, based upon your experience, you believe that the jury award will probably be greater than the settlement offer. Is there any question about what to do?

The decision about whether to accept the settlement offer could be guided by the following simple decision rule: When faced with uncertain outcomes, choose that decision with the highest expected value. However, such a decision rule often does not comport with the way

people actually make decisions. While the expected jury award is \$1,125,000, there is a 30% chance that the award will be \$750,000 — \$250,000 less than the settlement offer. Are you willing to advise your client to take this risk? Moreover, are there any professional responsibility implications for advising your client to take this risk? On the other hand, there is a 20% chance that the jury will award \$2,000,000, which exceeds the settlement value by \$1,000,000. There is the possibility of doubling your client's money by taking the risk.

The analysis in the preceding paragraph suggests that there might be some level of variability in the results that would discourage some individuals from choosing the option with the highest expected value. Greater variability implies that there is less certainty regarding the potential outcome, and thus more risk involved in the decision. Consider the same example, only this time, you feel that there is a 40% chance that the jury will find in favor of the defendant and award \$0, a 50% chance that the jury will award \$1,000,000, and a 10% chance that the jury will award \$5,250,000. Using the same formula as above, the expected value is still the same, even though there plainly is greater variability in outcomes.

$$E(x) = 0.4(\$0) + 0.5(\$1,000,000) + 0.1(\$6,250,000) = \$1,125,000$$

To the extent that individuals are concerned about the riskiness of decisions, they are interested in some measure of the variability of outcomes for any particular decision.

Variability, or risk, is quantified by a measure of the average dispersion of actual results around the expected value, which is known as the **variance**. Variance is equal to the weighted average of squared deviations from the expected value:

$$\sigma^2 = P_1(x_1 - E(x))^2 + P_2(x_2 - E(x))^2 + \dots + P_n(x_n - E(x))^2$$

where  $\sigma^2$  is the variance,  $P_i$  the probability of the outcome  $I$ ,  $x_i$  is the actual result for outcome  $I$ , and  $E(x)$  is the expected value as calculated above. Thus, for the jury award contemplated earlier, the variance is:

$$\begin{aligned} \sigma^2 &= 0.3(\$750,000 - \$1,125,000)^2 + 0.5(\$1,000,000 - \$1,125,000)^2 \\ &\quad + 0.2(\$2,000,000 - \$1,125,000)^2 \\ &= \$203,125,000,000 \end{aligned}$$

Because of the squared terms in the variance calculation, the answer is not in the same units as the inputs — in this case, dollars. As a result, it is sometimes difficult to interpret the meaning of the variance calculation. This difficulty can be resolved by using standard deviation. **Standard deviation** is also a measure of the dispersion around expected values, and it is measured in the same units as the original inputs. Standard deviation is simply the square root of the variance. For the jury award example, the standard deviation of possible awards is:

$$\sigma = \$450,694$$

The variance and standard deviation for the greater variability example above are:

$$\begin{aligned} \sigma^2 &= \$3,014,062,500,000 \\ \sigma &= \$1,736,106 \end{aligned}$$

The decision as to whether a dispersion or variability of \$450,694 is too risky is a matter of individual preference. The key to analyzing decisions under uncertainty is to recognize the fundamental economic tradeoff involved in these decisions. Specifically, how willing is a particular individual to sacrifice some amount of expected value for a reduction in risk? Ultimately, the decision regarding whether to take the certain settlement or the uncertain jury award depends upon the decision maker's subjective attitude toward risk.

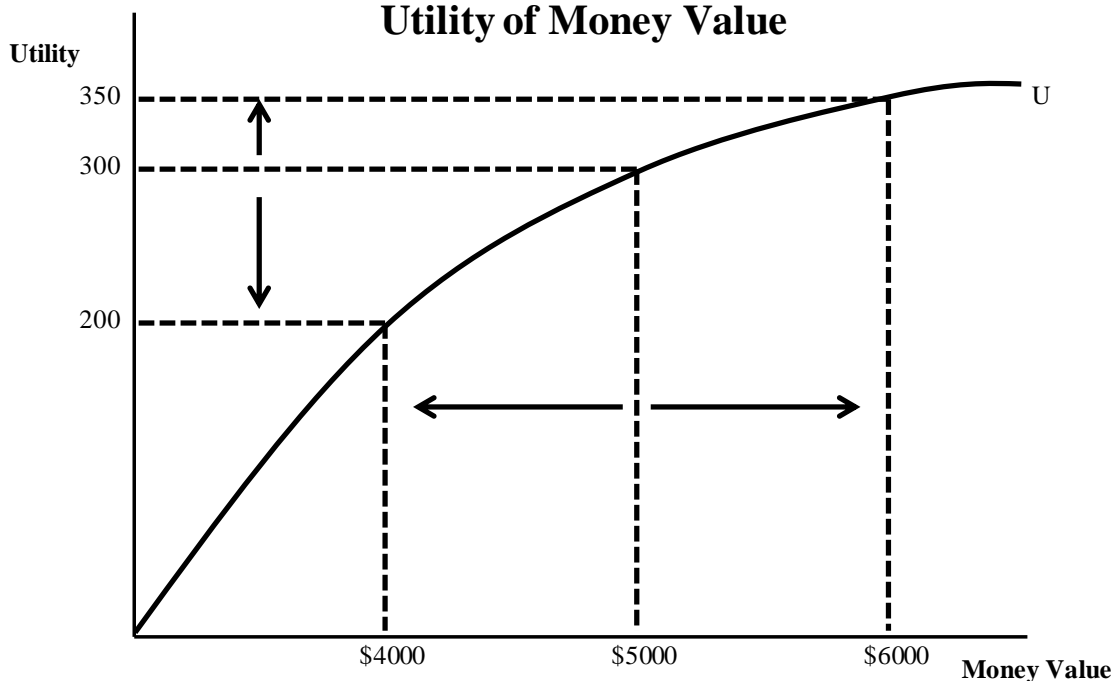
### 3. Expected Utility and Risk Preference

Most people do not like risk and thus might be willing to sacrifice some amount of

expected money value for a reduction in risk. People tend to prefer certain outcomes to uncertain outcomes, even if the expected value of the uncertain outcome is greater than that of the certain outcome. In other words, sometimes individuals get more utility — the benefit or satisfaction from a choice or course of action — from reducing risk than maximizing expected money outcomes. When faced with uncertainty, individuals attempt to maximize expected utility, not expected value.

Figure VI-1 presents a utility function for a hypothetical individual. Money value is marked along the horizontal axis and utility on the vertical axis — for this example, utility is a function of money value. Notice that the utility function,  $U$ , increases at a decreasing rate. This suggests that an individual with this type of utility function exhibits **diminishing marginal utility of money value**.

**Figure VI-1. Utility Curve Exhibiting Diminishing Marginal Utility of Money Value**



Consider an individual who starts off with \$5,000. Such an individual has a utility of 300 utils — an imaginary measurement of utility. If this individual's money holdings were increased by \$1,000, her utility would increase to 350 utils — an increase of 50 utils. On the other hand, if \$1,000 were taken away from this individual, her utility would decline to 200 utils — a loss of 100 utils. Thus, this individual's utility is affected more by losses than by gains: the loss of \$1,000 would decrease utility by a larger amount than utility increases when the individual gains \$1,000. This indicates that the individual exhibits diminishing marginal utility of money. Moreover, the individual's utility associated with a certain amount of \$5,000 is 300 utils. However, if she was involved in a lottery with a 50% chance of winning \$4,000 and a 50% chance of winning \$6,000, her expected utility would be only 275 utils ( $E(U)=0.5(200) + 0.5(350)$ ) even though the expected value of the lottery winnings was the same \$5,000. We can

conclude from this analysis that the individual is **risk averse**. Risk averse individuals prefer certain outcomes over uncertain outcomes that have the same expected value and exhibit diminishing marginal utility of money.

Figure VI-2 presents the dynamics of decision making under uncertainty for the risk averse individual. The uncertain decision faced by this individual has an expected value of \$10,000. However, there is a 50% chance that the payoff from this uncertain decision will be \$15,000 and a 50% chance of a \$5,000 payoff. The utility curve indicates how this individual regards uncertain decisions. Notice that a certain \$15,000 payoff provides the individual with utility corresponding to level A. A certain \$5,000 payoff provides utility that corresponds to level B. Furthermore, the utility from a certain \$10,000 payoff provides utility up to level C. However, the uncertain decision does not provide \$10,000 with certainty. In fact, although \$10,000 is the expected value, it will not occur. When the expected value is some combination of \$5,000 and \$15,000, the expected utility from the payoff can be determined by using the dotted line segment from E to F. The expected utility associated with an expected value of \$10,000 corresponds to level D — found by moving directly to the left from the dotted line segment until intersecting with the utility curve. The expected utility from an expected value of \$10,000 — level D — is less than the utility associated with a certain \$10,000 payoff — level C. In general, greater concavity in the utility function — the more the marginal utility of money value diminishes — suggests that individuals are more risk averse and less willing to accept uncertain outcomes.

**Figure VI-2. Decision Making under Uncertainty for Risk Averse Individuals**

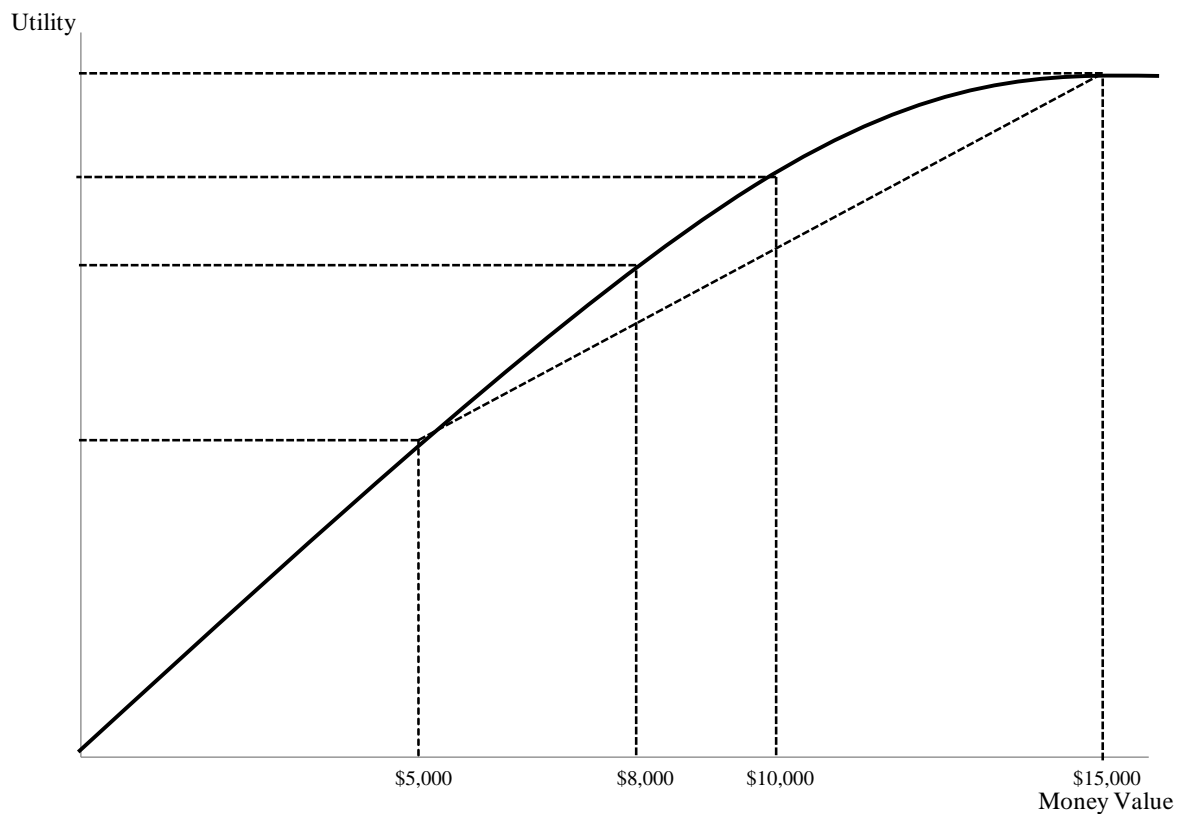


Figure VI-2 also demonstrates that risk averse individuals might be willing to sacrifice some amount of money value for reduced risk. Notice that a certain \$8,000 payoff provides the

same utility as the expected uncertain \$10,000 payoff. The \$8,000 payoff is known as the **certainty equivalent** of the expected uncertain \$10,000 payoff. The certainty equivalent of an uncertain amount is the sum of money which, if received with certainty will yield the same utility as the uncertain amount. Thus, risk averse individual might be willing to sacrifice an amount equal to \$2,000 ( $\$10,000 - \$8,000$ ) to avoid having to take the uncertain outcome. This \$2,000 payment is known as a **risk premium**. The risk premium is the maximum amount a person is prepared to pay to avoid an uncertain outcome. For example, this amount would correspond to how much an individual is willing to pay for fire insurance and avoid the risk of losing the entire value of her home if it burns down. This topic will be addressed in greater detail in the section on insurance.

Risk aversion describes the preferences revealed by much of the normal behavior observed everyday. Yet, this does not mean that all individuals are risk averse. Some individuals are **risk neutral** — they appear to be indifferent toward risk. Moreover, some individuals seem to actively seek risky situations. Such individuals are known as **risk seekers**. Both risk neutral and risk seeking behavior can be incorporated into expected utility analysis.

Figure VI-3 presents expected utility analysis for the risk neutral individual. Notice that, unlike the utility curve of the risk averse individual, which exhibits diminishing marginal utility of money, the utility curve of the risk neutral individual increases at a constant rate. This suggests that an increase in money holdings would have the same marginal effect on utility as an equal loss in money holdings. Consider the relationship between the utility for a certain payoff of \$10,000 and an uncertain expected payout of \$10,000. The straight utility function indicates that the utility of the certain outcome is the same as the utility for the uncertain outcome. Thus, risk neutral individuals are indifferent between certain and uncertain outcomes — their decisions are guided only by expected value. To these individuals, there is no difference between a \$1 million bet and a \$1 bet so long as the expected value is the same. When risk neutral individuals maximize their expected value, they are also maximizing their expected utility.

Figure VI-2. Decision Making under Uncertainty for Risk Averse Individuals

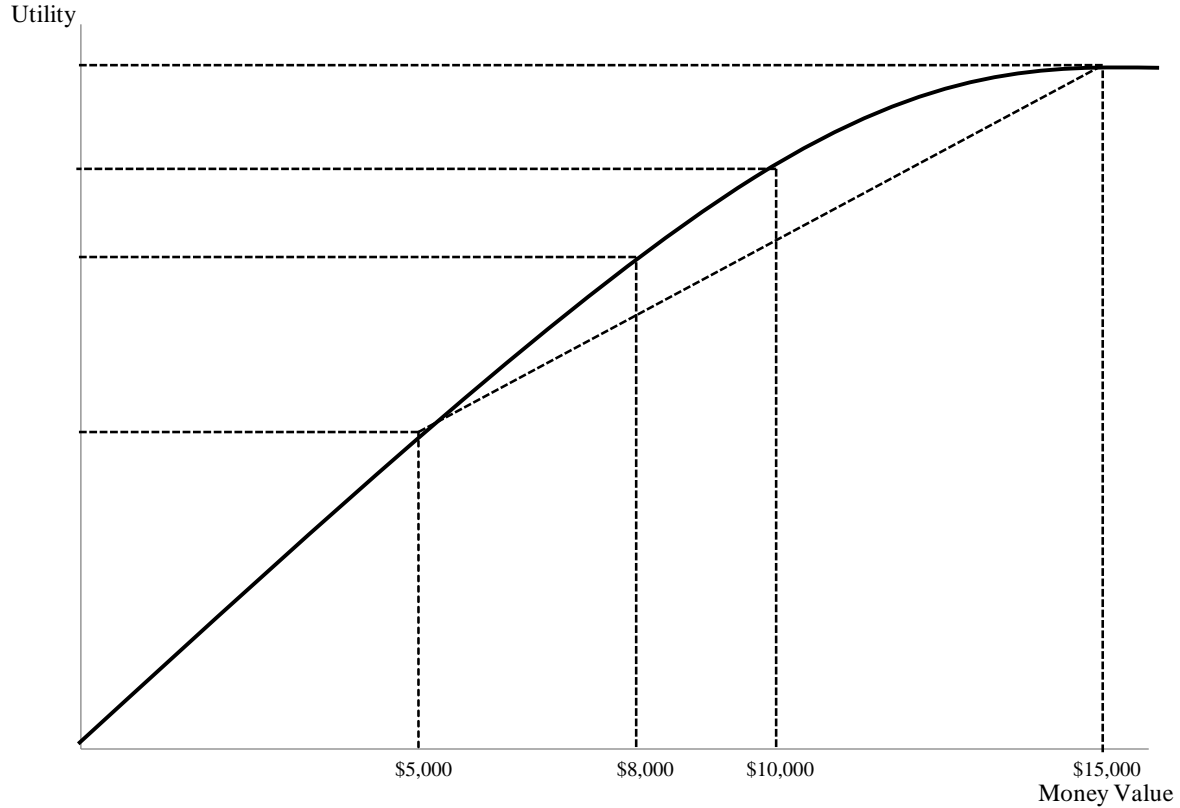
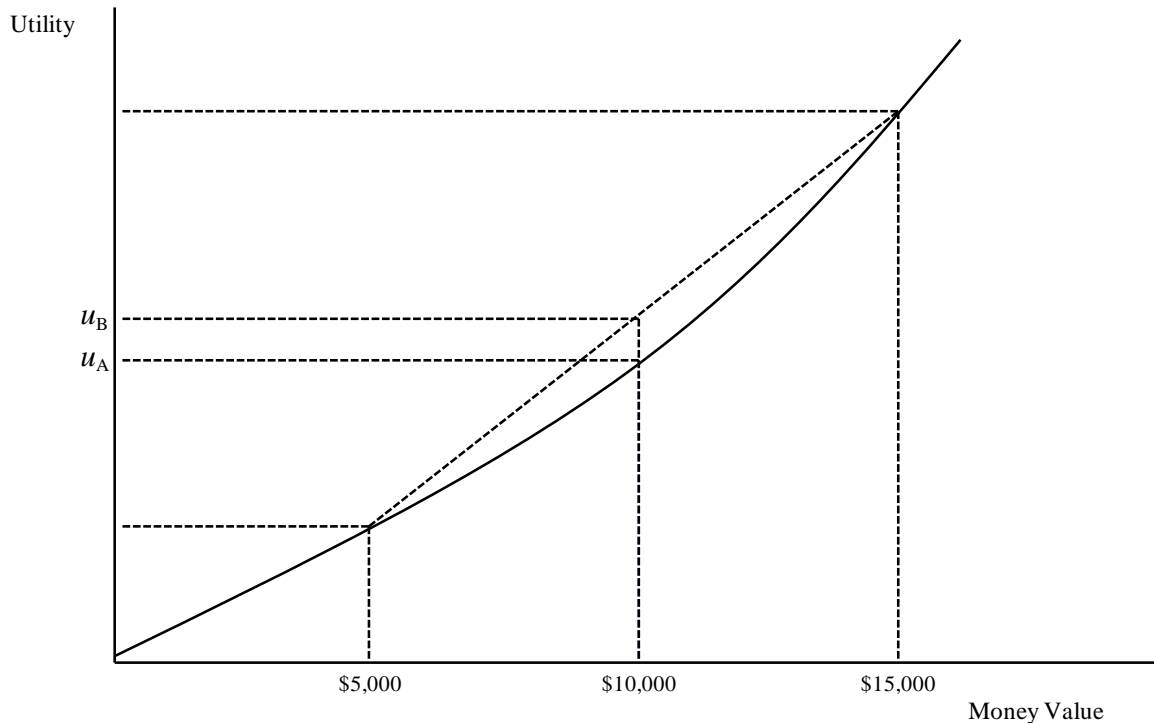


Figure VI-4 presents the expected utility analysis for a risk-seeking individual. Notice that the utility curve increases at an increasing rate, exhibiting increasing marginal utility of money. This suggests that an increase in money holdings would have a greater marginal effect on utility than would an equal loss in money holdings. Consider the relationship between the utility for a certain payoff of \$10,000 and an uncertain expected payout of \$10,000. Because of the upward sloping nature of the utility function, the utility of the certain outcome ( $u_A$ ) is less than the utility for the uncertain outcome ( $u_B$ ). Risk seekers prefer uncertain outcomes to certain outcomes.

**Figure VI-4. Risk Seeker**



A final possibility is suggested by what is called "prospect theory" — a theory developed by behavioral psychologists Daniel Kahneman and Amos Tversky (recall the discussion of behavioral law and economics in Part A.2 of Chapter I). As explained by Professor Chris Guthrie, one of the central components of prospect theory is that

... people evaluate decision options relative to some reference point, generally the status quo. When choosing between options that appear to be gains relative to that reference point, people tend to make risk-averse choices; when choosing between options that appear to be losses, people tend to make risk-seeking choices. For example, people will generally choose a definite \$1,000 prize over a 50% chance at receiving a \$2,000 prize but will opt to face a 50% chance at having to pay a \$2,000 fine over having to pay a definite \$1,000 fine. This result is inconsistent with rational choice theory, which generally assumes either risk neutrality or risk aversion in the face of both gains and losses.

Chris Guthrie, *Prospect Theory, Risk Preference, and the Law*, 97 *Northwestern University Law Review* 1115, 1118–19 (2003). In other words, prospect theory holds that people's risk preferences vary. For events that occur with moderate or high probability, most people are risk averse when facing the prospect of a gain, but are risk seeking when facing a loss. However, for events that occur with low probability, the reverse is true: most people are risk seeking with gains (they play the lottery) and risk averse with losses (they buy insurance).

How might the predictions from prospect theory be used to explain the unwillingness of plaintiffs (who face the prospect of a gain) and defendants (who face the prospect of a loss) to



settle cases with a positive settlement range?

## **B. Insurance**

People are generally risk averse concerning risks that potentially affect a significant proportion of their wealth. An important behavioral implication of such risk aversion is that these individuals are willing to sacrifice certain amounts of money (up to their risk premium) in order to avoid the prospect of large losses. An example of this is the decision to purchase insurance. People who purchase health insurance, automobile insurance, disability insurance, or homeowners insurance give up a certain relatively small sum of money in order to avoid an uncertain, yet potentially larger, loss. This section considers the demand and supply of insurance, as well as some of the problems facing insurance providers.

### **1. Demand for Insurance**

The demand for insurance is derived from individuals' risk preferences. Perhaps the best way to see this is to consider an everyday example. Suppose an individual owns a home that is worth \$100,000. If there is a one percent chance that the home will be destroyed by fire in the next year, the expected loss from the fire is \$1,000. If an individual does not purchase insurance, one of two things will happen: the house will survive the year unscathed and there will be no loss; or, the house will be destroyed by fire and there will be a \$100,000 loss. Thus, the expected loss for the year is \$1,000, although the individual will never be exactly \$1,000 dollars worse off. The individual will either lose nothing or lose \$100,000. Intuitively, the variance across possible outcomes is wide. A risk averse individual might be willing to pay, let's say, up to \$2,000 for insurance to guarantee that the house is replaced in the event that it is destroyed. The individual's position for the year then is \$98,000 — \$100,000 minus the \$2,000 insurance premium — with certainty as opposed to an uncertain \$99,000 — \$100,000 minus the expected loss of \$1,000. As a result of purchasing insurance, there is only one possible outcome — \$98,000 — and the variance of possible outcomes is zero. In this case, insurance reduces the individual's financial risk to zero.

### **2. Supply of Insurance**

Insurance companies are in a much different situation than the individual homeowner. Insurance companies are in the business of pooling the risk of a large number of homeowners. A statistical phenomenon known as the "law of large numbers" allows insurance companies to pool the risks of a large number of homeowners. The law of large numbers holds that as the pool size increases, unpredictable events become more predictable. For example, an insurance company can estimate with near certainty the number of homes that will be destroyed by fire. If all homes have a one percent chance of being destroyed by fire in the next year, and the insurance company has written policies for 1,000 homes, then the expected number of homes to burn in any year is one percent times 1,000 = 10. Assuming an average loss of \$100,000 per home, the insurance company would expect to pay out \$1,000,000 per year for 10 houses burning down. Due to the law of large numbers, the variance of the insurance company's estimates approaches certainty as more homes are added to the risk pool. Moreover, as the number of insured homes increases, the insurance company becomes indifferent to the risk or loss; it deals purely with the monetary calculations.

However, each individual homeowner is still faced with the large potential variance in the results between losing \$0 or \$100,000. The expected claim per policy for the insurance company

is \$1,000 — that is, the \$1,000,000 the insurer expects to pay out divided by the 1,000 homes insured. Earlier, it was assumed that homeowners are willing to pay up to \$2,000 per insurance policy. The difference between the price the consumer is willing to pay and the cost to the insurance company of providing the insurance creates an opportunity for an insurance market to emerge.

Insurance companies also have administrative costs associated with the selling and pricing of policies, servicing policy holders, and adjusting claims. Successful insurance companies must price their policies so they cover their expected losses and their administrative costs. Moreover, insurance companies that do a better job of investing their cash reserves (before they have to pay out the claims owed to policyholders) receive a larger return and are able to offer lower prices. To the extent that the insurance market is competitive, prices are pushed below the \$2,000 per policy that consumers are willing to pay towards a price that reflects the cost of operating in the industry.

Controlling the cost of operation is an important challenge for all insurance companies. Every insurer must deal with two basic problems — adverse selection and moral hazard.

#### *a. Adverse Selection*

**Adverse selection** is a precontractual problem that derives from the fact that potential insured parties know much more about their own risk than does the potential insurer. This information is asymmetric, meaning it is difficult for the insurer to determine the risk characteristics of the potential insured. As a result of this asymmetric information, people who are risky (that is, pose a greater than average risk) are more likely to attempt to purchase insurance than are people who have a below average risk. This adverse selection can lead to a "lemons market" — a concept that was introduced in Chapter V — a phenomenon that has the potential to destroy markets whenever there is asymmetric information.

Insurance companies attempt to deal with adverse selection by providing exclusions for coverage dealing with preexisting conditions, looking for various statistical phenomena that are likely to result in higher claims from various groups, checking out applicants' backgrounds, and a variety of other devices that are designed to avoid ending up with a pool of lemons. Sometimes risky insureds might be allowed into the pool, but they will have to pay higher rates. Typical examples of ways to deal with this type of adverse selection include charging higher homeowners and health insurance rates for people who smoke, higher automobile insurance rates for males under age 25, and so forth.

### **Hall v. Continental Casualty Co.**

United States District Court for the Western District of Wisconsin  
207 F. Supp. 2d 903 (2002)

BARBARA B. CRABB, District Judge.

In this civil action for monetary relief, plaintiff Valerie K. Hall contends that defendant Continental Casualty Company breached its long-term disability insurance policy by denying her claim for long-term disability benefits on the ground that plaintiff's lung cancer is a pre-existing condition excluded from coverage. Plaintiff also alleges that defendant denied her claim for benefits in bad faith. . . .

\* \* \*

OPINION

A. Breach of Contract

It is common for health and long-term disability insurance policies to exclude the treatment of pre-existing conditions from coverage. Although the definition of "pre-existing condition" can vary, the policy reasons for enforcing pre-existing condition exclusions are uniform. First, losses stemming from pre-existing conditions are not fortuitous. In the case of a true pre-existing condition, the insured knows that she will be incurring losses because of the condition. Second, losses stemming from pre-existing conditions have the potential of undermining the insurer's risk pool. If an insured does not disclose to the insurer that she has a pre-existing condition when applying for coverage, the insurer cannot calculate its potential liability accurately. Finally, such losses create problems of adverse selection. An insured who has been diagnosed with a condition could apply for insurance without revealing the condition, costing the insurer more than it had anticipated.

Although these policy concerns are economically sound, extending pre-existing condition exclusions too far has a deleterious effect on the insured. If undergoing a standard diagnostic test before the effective date of a policy could be considered receiving treatment for a pre-existing condition, individuals could be discouraged from seeking preventive medical care. If an insurer can point to pre-coverage symptoms that turn out to be consistent with a condition diagnosed after coverage becomes effective, any prior symptoms not inconsistent with the ultimate diagnosis could become a ground for denying benefits. *Ermenc v. American Family Mutual Ins. Co.*, 221 Wis. 2d 478, 484, 585 N.W.2d 679, 682 (Ct. App. 1998). Similarly, if an insurer can point to the fact that doctors suspected a condition before an effective date because of certain risk factors, such as a history of smoking, the definition of pre-existing condition would become so broad as to make the term meaningless. For example, it is likely that any doctor would suspect cancer as a possible cause of respiratory problems in a heavy smoker. It would be absurd to find this suspicion sufficient to deny coverage of all subsequent treatment on the ground that cancer is a pre-existing condition. Because "insurance contracts should be given a reasonable interpretation and not one that leads to an absurd result," these policy concerns demand that pre-existing condition exclusions not be read too broadly.

Under the terms of plaintiff's policy in this case, a pre-existing condition is "a condition for which medical treatment or advice was rendered, prescribed or recommended within three months prior to [the insured's] effective date of insurance." The central issue is whether plaintiff's lung cancer constitutes a pre-existing condition under plaintiff's policy, which both parties agree became effective on September 1, 1999. . . .

. . . In *Ermenc*, the leading Wisconsin case on the coverage of pre-existing condition exclusions, the Wisconsin Court of Appeals held that the insured's cancer was not a pre-existing condition despite pre-existing symptoms. Although defendant tries to distinguish *Ermenc*, I find the reasoning dispositive.

\* \* \*

The facts of this case are similar to those in *Ermenc*. Before the effective date of her policy, plaintiff went to the emergency room complaining of chest pain, shortness of breath and numbness in her left arm and was diagnosed as having pneumonia, tachycardia and asthma. X-rays taken that day revealed a spot on her lung that could have resulted from several conditions, including pneumonia or cancer. The radiologist who interpreted the x-ray recommended a follow-up x-ray to determine whether the spot was caused by pneumonia. If not, further testing would be necessary to determine the cause of the spot and to rule out cancer. After the effective date of plaintiff's insurance policy, plaintiff's primary physician ordered follow-up x-rays and tests that led to a diagnosis of lung cancer.

The only material difference between the facts in this case and those in *Ermenc* is that here, plaintiff's doctors suspected cancer before the effective date. Defendant relies heavily on this fact, asserting that the fact that Ermenc's doctors did not suspect cancer led the court to conclude that they did not give her advice about cancer or treat her for it. However, in neither case did the doctors make a diagnosis of cancer before the effective date. In plaintiff's case, the doctors narrowed the possible diagnoses to a handful of conditions, one of which was cancer; they did not order the follow-up x-rays solely to rule out the possibility of cancer. Moreover, the facts suggest that even if the emergency room doctors had not suspected cancer, they would have ordered the same sequence of x-rays in order to track plaintiff's pneumonia. Because plaintiff had been a smoker for many years, she had a high risk factor for lung cancer, making it an obvious condition to suspect. As in *Ermenc*, plaintiff's symptoms that manifested before the effective date were non-specific and could have been caused by a variety of conditions. Defendant has not shown that plaintiff was treated for lung cancer both before and after the effective date of plaintiff's policy, or that she was given advice about cancer before the effective date. Taking all facts in the light most favorable to the non-moving party, I conclude that defendant has failed to establish that plaintiff's cancer was a condition for which "medical treatment or advice was rendered, prescribed or recommended" in the three months before the policy took effect.

Policy considerations reinforce the conclusion that plaintiff did not receive treatment or advice for cancer before the effective date of her policy. None of the concerns that weigh in favor of enforcing pre-existing condition clauses come into play. Neither plaintiff nor her doctors knew that she had cancer before the effective date of her policy; she did not discover that she had cancer and then purchase a long-term disability policy from defendant. Although the cost of providing long-term disability insurance to plaintiff may cost defendant more than it had anticipated, this circumstance is not through any fault of plaintiff. It falls instead to the insurer and its risk pool.

In contrast, policy considerations that weigh in favor of protecting the insured do come into play. When plaintiff underwent a chest x-ray at the emergency room and when the emergency room doctors recommended a follow-up x-ray, they were following standard diagnostic procedures, which should not be discouraged. Although the doctors suspected that the spot on plaintiff's lung could result from cancer, this suspicion was based, in part, on plaintiff's history as a smoker, which put her at greater risk for lung cancer. Individuals should not be denied coverage because they carry risk factors that make them more prone to certain health conditions that, in turn, make a doctor more suspicious of these ailments. Instead, insurers should (and do) consider such risk factors when computing insurance premiums.

\* \* \*

#### B. Bad Faith Denial

To prevail on a claim for bad faith, an insured must establish that (1) there was no reasonable basis for denying the claim under an objective standard and (2) the insurer acted with knowledge or reckless disregard for the lack of a reasonable basis. . . . When an insured's claim is "fairly debatable" either in fact or law, an insurer cannot be said to have denied the claim in bad faith. . . . [B]ecause I find that the law addressing pre-existing condition clauses is "fairly debatable," defendant's denial of coverage cannot be characterized as a bad faith denial.

Although the undisputed facts are insufficient to establish that defendant investigated plaintiff's claim reasonably, this lack of evidence is not fatal to defendant's argument. On the basis of the undisputed facts surrounding plaintiff's medical treatment as applied to case law, it is "fairly debatable" whether plaintiff's lung cancer constituted a pre-existing condition. Although I

am not convinced by the factual distinctions that defendant tries to draw between the facts of this case and those in *Ermenc*, there is room to argue that the two cases are distinguishable. In addition, the survey of federal and state court cases that the parties undertook demonstrates that other jurisdictions have reached different conclusions on the basis of arguably similar facts. Because there is an arguable basis for distinguishing *Ermenc* and other jurisdictions have found pre-existing condition clauses to apply in relatively similar circumstances, I conclude that plaintiff's claim for coverage was "fairly debatable," precluding a finding that defendant denied coverage in bad faith. Defendant's motion for summary judgment as to this claim will be granted.

\* \* \*

### Notes and Questions

**1. Adverse Selection:** The selection of people who purchase insurance is not a normal, random sample of the population, but rather includes some people with private information about their personal situations that makes it likely they will receive a higher than average level of benefit under the insurance policy. For example, high-risk patients (e.g., cancer-prone, drug users, etc.) apply to insurance companies, while young healthy persons do not. Insurance deals with this adverse selection problem in a variety of ways. The most obvious is to require the disclosure of information on the long application forms with boiler-plate provisions to try and keep out risky patients. Insurance companies set limits on pre-existing conditions, through either limited coverage or higher premiums for the additional risk.

**2. Regulation by the Court:** Isn't the court saying that whenever there is a discrepancy between the insurance company and the consumer, the consumer gets the benefit of the doubt if he or she acted in good faith? If so, isn't the court simply regulating the market? Companies will increase premiums on all consumers to pay for those consumers who "slip through the cracks." Thus, we all pay for this ability to receive the benefit of the doubt.

**3. Adverse Selection and the Affordable Care Act:** The existence of adverse selection in the market for health insurance was one of the justifications for the "individual mandate" of the Patient Protection and Affordable Care Act of 2010, which requires individuals to buy a certain minimum amount of health insurance coverage or face a penalty. The Supreme Court described the issue in its decision upholding the Affordable Care Act as follows:

The Government's first argument is that the individual mandate is a valid exercise of Congress's power under the Commerce Clause and the Necessary and Proper Clause. According to the Government, the health care market is characterized by a significant cost-shifting problem. Everyone will eventually need health care at a time and to an extent they cannot predict, but if they do not have insurance, they often will not be able to pay for it. Because state and federal laws nonetheless require hospitals to provide a certain degree of care to individuals without regard to their ability to pay, see, e.g., [42 U. S. C. §1395dd](#); Fla. Stat. Ann. §395.1041, hospitals end up receiving compensation for only a portion of the services they provide. To recoup the losses, hospitals pass on the cost to insurers through higher rates, and insurers, in turn, pass on the cost to policy holders in the form of higher premiums. Congress estimated that the cost of uncompensated care raises family health insurance premiums, on average, by over \$1,000 per year.

In the Affordable Care Act, Congress addressed the problem of those who cannot obtain insurance coverage because of preexisting conditions or other health issues. It did so through the Act's "guaranteed-issue" and "community-rating" provisions. These provisions together prohibit insurance companies from denying coverage to those with

such conditions or charging unhealthy individuals higher premiums than healthy individuals.

The guaranteed-issue and community-rating reforms do not, however, address the issue of healthy individuals who choose not to purchase insurance to cover potential health care needs. In fact, the reforms sharply exacerbate that problem, by providing an incentive for individuals to delay purchasing health insurance until they become sick, relying on the promise of guaranteed and affordable coverage. The reforms also threaten to impose massive new costs on insurers, who are required to accept unhealthy individuals but prohibited from charging them rates necessary to pay for their coverage. This will lead insurers to significantly increase premiums on everyone.

The individual mandate was Congress's solution to these problems. By requiring that individuals purchase health insurance, the mandate prevents cost-shifting by those who would otherwise go without it. In addition, the mandate forces into the insurance risk pool more healthy individuals, whose premiums on average will be higher than their health care expenses. This allows insurers to subsidize the costs of covering the unhealthy individuals the reforms require them to accept. The Government claims that Congress has power under the Commerce and Necessary and Proper Clauses to enact this solution.

National Federation of Independent Business v. Sebelius, 132 S. Ct. 2566, 2585 (2012).

#### ***b. Moral Hazard***

Moral hazard is the problem that arises when the insured's behavior changes after purchasing insurance, so that the probability of loss or size of loss increases. Moral hazard is a post-contractual problem concerned with insured parties changing their behavior once they have become insured. Insured individuals change their behavior because the risk has now been shifted to the insurer. Thus, the insured no longer have the same incentive to take risk-reducing precautions. The changes in behavior can range from an individual deciding to take up skydiving shortly after purchasing a life insurance policy to exercising less care with respect to preventing a car or other personal property from being stolen when it is protected by insurance.

Insurance companies must recognize these moral hazard incentives and develop methods to deal with them. Insurance companies deal with the moral hazard problem by a combination of deductibles and copayments. Deductibles are the amount that must be paid by the insured party prior to any payment by the insurer; insured parties essentially "self-insure" up to the amount of the deductible. Once the deductible is met, the insurer pays the rest. By requiring the insured to pay a portion of any loss, deductibles give the insured an incentive to avoid the loss, and thus reduce the moral hazard problem. Copayments are payments by the insured of a proportion of any loss; the insurer pays the remaining share. Copayments similarly deter moral hazard by forcing insured parties to pay for at least a portion of the consequences of their risky behavior.

**Atwater Creamery Co. v.  
Western National Mutual Insurance Co.**  
Supreme Court of Minnesota  
366 N.W.2d 271 (1985)

WAHL, Justice.

Atwater Creamery Company (Atwater) sought a declaratory judgment against its insurer, Western National Mutual Insurance Company (Western), seeking coverage for losses sustained during a burglary of the creamery's storage building. . . . The Kandiyohi County District Court . .

. dismissed the jury for lack of disputed issues of fact and ordered judgment in favor of the insurer, concluding that the burglary insurance policy in effect defined burglary so as to exclude coverage of this burglary. We . . . reverse as to the policy coverage.

Atwater does business as a creamery and as a supplier of farm chemicals in Atwater, Minnesota. It was insured during the time in question against burglary, up to a ceiling of \$20,000, by Western under Mercantile Open Stock Burglary Policy SC10-1010-12, which contained an "evidence of forcible entry" requirement in its definition of burglary. The creamery had recovered small amounts under this policy for two separate burglaries prior to the events in this case.

Atwater built a separate facility, called the Soil Center, a few blocks away from its main plant in 1975 for the purpose of storing and selling chemicals. The Soil Center is a large rectangular building with two regular doors along the north side and two large, sliding doors, one each on the east and west sides. There are no other entrances or exits to or from the building itself. One of the doors on the north side leads into the office in the northwest corner of the building. It is secured by a regular dead bolt lock, opened with a key. There is no access into the main portion of the building from the office. Persons entering the main area must use the other door on the north side which is secured by a padlock after hours. The large sliding doors on the east and west are secured by large hasps on each side of each door which are held tight by turnbuckles that must be loosened before the doors can be opened.

Inside the main area of the building, along the north wall, is a large storage bin with three separate doors, each of which is secured by a padlock. Between the storage bin and the office is an "alleyway," entered through the large sliding doors, which runs east and west the length of the building. Trucks are stored in the alleyway when not in use.

Sometime between 9:30 p.m., Saturday, April 9, and 6 a.m., Monday, April 11, 1977, one or more persons made unauthorized entry into the building, took chemicals worth \$15,587.40, apparently loading them on the truck that had been parked inside and driving away after loosening the turnbuckles on the east door and closing it. The truck was later found parked near the town dump, with the key still in the ignition.

Larry Poe, the plant manager at the Soil Center, had left at 9:30 p.m. on Saturday, after making sure everything was properly secured. On Monday morning, the north side doors were locked securely, but two of the three doors to the storage bin were ajar. Their padlocks were gone and never found. The turnbuckles had been loosened on the east sliding door so that it could be easily opened or closed.

An investigation by the local police, the Kandiyohi County Sheriff's Department, and the Minnesota Bureau of Criminal Investigation determined that no Atwater Creamery employees, past or present, were involved in the burglary. Suspicion settled on persons wholly unconnected with the creamery or even with the local area, but no one has been apprehended or charged with the crime.

Atwater filed a claim with Western under the burglary policy. Western denied coverage because there were no visible marks of physical damage to the exterior at the point of entrance or to the interior at the point of exit, as required by the definition of burglary in the policy. The creamery then brought suit against Western for the \$15,587.40 loss, \$7,500 in other directly related business losses and costs, disbursements and reasonable attorney fees.

Charles H. Strehlow, the owner of the Strehlow Insurance Agency in Willmar, Minnesota, and Western's agent, testified that he is certain he mentioned the evidence-of-forcible-entry requirement to Poe and members of the Atwater Board of Directors but was unable

to say when the discussion occurred. Poe and the board members examined do not remember any such discussion. None of the board members had read the policy, which is kept in the safe at the main plant, and Poe had not read it in its entirety. He stated that he started to read it but gave up because he could not understand it.

The issue[] on appeal [is] . . . whether the reasonable expectations of the insured as to coverage govern to defeat the literal language of the policy. . . .

\* \* \*

The definition of burglary in this policy is one used generally in burglary insurance. Courts have construed it in different ways. It has been held ambiguous and construed in favor of coverage in the absence of visible marks of forcible entry or exit. We reject this analysis because we view the definition in the policy as clear and precise. It is not ambiguous.

In determining the intent of the parties to the insurance contract, courts have looked to the purpose of the visible-marks-of-forcible-entry requirement. These purposes are two: to protect insurance companies from fraud by way of "inside jobs" and to encourage insureds to reasonably secure the premises. As long as the theft involved clearly neither an inside job nor the result of a lack of secured premises, some courts have simply held that the definition does not apply.

In the instant case, there is no dispute as to whether Atwater is attempting to defraud Western or whether the Soil Center was properly secured. The trial court found that the premises were secured before the robbery and that the law enforcement investigators had determined that it was not an "inside job." To enforce the burglary definition literally against the creamery will in no way effectuate either purpose behind the restrictive definition. We are uncomfortable, however, with this analysis given the right of an insurer to limit the risk against which it will indemnify insureds.

At least three state courts have held that the definition merely provides for one form of evidence which may be used to prove a burglary and that, consequently, other evidence of a burglary will suffice to provide coverage. The Nebraska Supreme Court recently rejected this argument[, holding] . . . that the definition is not a rule of evidence but is a limit on liability, is unambiguous and is applied literally to the facts of the case at hand. We, too, reject this view of the definition as merely a form of evidence. The policy attempts to comprehensively define burglaries that are covered by it. In essence, this approach ignores the policy definition altogether and substitutes the court's or the statute's definition of burglary. This we decline to do, either via the conformity clause or by calling the policy definition merely one form of evidence of a burglary.

Some courts and commentators have recognized that the burglary definition at issue in this case constitutes a rather hidden exclusion from coverage. Exclusions in insurance contracts are read narrowly against the insurer. Running through the many court opinions refusing to literally enforce this burglary definition is the concept that the definition is surprisingly restrictive, that no one purchasing something called burglary insurance would expect coverage to exclude skilled burglaries that leave no visible marks of forcible entry or exit. Professor Robert E. Keeton, in analyzing these and other insurance cases where the results often do not follow from the rules stated, found there to be two general principles underlying many decisions. These principles are the reasonable expectations of the insured and the unconscionability of the clause itself or as applied to the facts of a specific case. . . .

The doctrine of protecting the reasonable expectations of the insured is closely related to the doctrine of contracts of adhesion. Where there is unequal bargaining power between the parties so that one party controls all of the terms and offers the contract on a take-it-or-leave-it



basis, the contract will be strictly construed against the party who drafted it. Most courts recognize the great disparity in bargaining power between insurance companies and those who seek insurance. Further, they recognize that, in the majority of cases, a lay person lacks the necessary skills to read and understand insurance policies, which are typically long, set out in very small type and written from a legalistic or insurance expert's perspective. Finally, courts recognize that people purchase insurance relying on others, the agent or company, to provide a policy that meets their needs. The result of the lack of insurance expertise on the part of insureds and the recognized marketing techniques of insurance companies is that "[t]he objectively reasonable expectations of applicants and intended beneficiaries regarding the terms of insurance contracts will be honored even though painstaking study of the policy provisions would have negated those expectations."

The traditional approach to construction of insurance contracts is to require some kind of ambiguity in the policy before applying the doctrine of reasonable expectations. Several courts, however, have adopted Keeton's view that ambiguity ought not be a condition precedent to the application of the reasonable-expectations doctrine.

As of 1980, approximately ten states had adopted the newer rule of reasonable expectations regardless of ambiguity. Other states, such as Missouri and North Dakota, have joined the ten since then. Most courts recognize that insureds seldom see the policy until the premium is paid, and even if they try to read it, they do not comprehend it. Few courts require insureds to have minutely examined the policy before relying on the terms they expect it to have and for which they have paid.

The burglary definition is a classic example of a policy provision that should be, and has been, interpreted according to the reasonable expectations of the insured. . . .

Atwater had a burglary policy with Western for more than 30 years. The creamery relied on Charles Strehlow to procure for it insurance suitable for its needs. There is some factual dispute as to whether Strehlow ever told Poe about the "exclusion," as Strehlow called it. Even if he had said that there was a visible-marks-of-forcible-entry requirement, Poe could reasonably have thought that it meant that there must be clear evidence of a burglary. There are, of course, fidelity bonds which cover employee theft. The creamery had such a policy covering director and manager theft. The fidelity company, however, does not undertake to insure against the risk of third-party burglaries. A business that requests and purchases burglary insurance reasonably is seeking coverage for loss from third-party burglaries whether a break-in is accomplished by an inept burglar or by a highly skilled burglar. Two other burglaries had occurred at the Soil Center, for which Atwater had received insurance proceeds under the policy. Poe and the board of the creamery could reasonably have expected the burglary policy to cover this burglary where the police, as well as the trial court, found that it was an "outside job."

The reasonable-expectations doctrine gives the court a standard by which to construe insurance contracts without having to rely on arbitrary rules which do not reflect real-life situations and without having to bend and stretch those rules to do justice in individual cases. As Professor Keeton points out, ambiguity in the language of the contract is not irrelevant under this standard but becomes a factor in determining the reasonable expectations of the insured, along with such factors as whether the insured was told of important, but obscure, conditions or exclusions and whether the particular provision in the contract at issue is an item known by the public generally. The doctrine does not automatically remove from the insured a responsibility to read the policy. It does, however, recognize that in certain instances, such as where major exclusions are hidden in the definitions section, the insured should be held only to reasonable

knowledge of the literal terms and conditions. The insured may show what actual expectations he or she had, but the factfinder should determine whether those expectations were reasonable under the circumstances.

\* \* \*

In our view, the reasonable-expectations doctrine does not automatically mandate either pro-insurer or pro-insured results. It does place a burden on insurance companies to communicate coverage and exclusions of policies accurately and clearly. It does require that expectations of coverage by the insured be reasonable under the circumstances. Neither of those requirements seems overly burdensome. Properly used, the doctrine will result in coverage in some cases and in no coverage in others.

We hold that where the technical definition of burglary in a burglary insurance policy is, in effect, an exclusion from coverage, it will not be interpreted so as to defeat the reasonable expectations of the purchaser of the policy. Under the facts and circumstances of this case, Atwater reasonably expected that its burglary insurance policy with Western would cover the burglary that occurred. Our holding requires reversal as to policy coverage.

### **Notes and Questions**

**1. *Why the Physical Damage Requirement?:*** There are two stated reasons for the physical damage requirement. First, insurance companies are trying to protect themselves from "inside jobs" by giving insured employers incentives to monitor their employees' fidelity. Clearly, the employers are in a better position to assess the risk of employee infidelity. Second, the requirement gives the insured company the incentive to invest in security precautions up to the point that any burglar must leave some physical trace of entry. Thus, the insurance company, by providing this "loophole," helps to align the insured's incentives to invest in the company's security.

**2. *Dealing with Moral Hazard:*** Insurance companies recognize the presence of moral hazard incentives and attempt to develop methods to deal with them. Clearly, Western's physical damage requirement attempts to address such issues. Given the outcome in *Atwater*, how might Western now go about protecting itself from moral hazard incentives? The court suggests that insurance companies now have a burden to communicate coverage and exclusions accurately and clearly. How would an insurance company prove that it did so? Would the insurance company face further incentive problems in collecting such evidence? One possibility is to have the insured place her initials on the contract next to the exclusions as the insurer explains them. Would such a practice help insurers to avoid the reasonable-expectations doctrine? Wouldn't the difficulty of obtaining proof of accurate and clear communication result in higher premiums, therefore excluding some from insurance altogether?

**3. *Getting What You Paid For:*** The general policy of construing exclusions narrowly makes it more difficult for insurance companies to protect against moral hazard incentives. In effect, such a policy expands the insurers' risk exposure. Atwater paid premiums that were related to Western's expected risk exposure. Who bears the cost when Western is forced to pay on a risk not contemplated by the insurance contract? Did Atwater get more than it paid for? Economic efficiency would seem to suggest that the least cost avoider of risk should bear the cost. Does the court's result achieve this end?

### ***c. Insurer's Duty to Settle***

The insurance market, like any other voluntary market, is founded on mutually beneficial

exchange. The claims payment process, as indicated by the prior cases in this section, can often result in conflicts between the insured and the insurer. Potential conflicts arise when a plaintiff is attempting to negotiate a settlement with an insured defendant and the defendant's insurance company. Consider the incentives faced by the insured and insurer in the following case.

**Mowry v. Badger State Mutual Casualty Co.**

Supreme Court of Wisconsin

385 N.W.2d 171 (1986)

CECI, J.

This is a review of the circuit court's decision and judgment against Badger State Mutual Casualty Company (Badger State) in the amount of \$159,000. In a decision filed on May 23, 1984, the circuit court for Waukesha county, Robert T. McGraw, circuit judge, held that Badger State breached its contract and committed bad faith in refusing to defend its insured and in refusing to settle the third-party claim of victim Bradley Mowry within the liability limits of an insurance policy. We reverse the judgment of the circuit court.

\* \* \*

The issue is whether the circuit court erred in holding that Badger State breached its contract with its insured and committed the tort of bad faith in refusing to defend its insured and in refusing to negotiate a settlement within policy limits when Badger State had sought a separate trial on the issue of coverage. . . .

The historical facts of this case are undisputed. On May 3, 1975, Bradley Mowry, then age 19, was injured in an automobile accident and suffered serious bodily injury, including the amputation of part of one foot. He was a passenger in an automobile driven by Steven McCarthy. The vehicle left a roadway and collided with a bridge abutment. McCarthy's parents were insured by Badger State and had policy limits of \$15,000 for damages to any one person and medical coverage up to \$1,000.

Upon being notified of the accident, Badger State began to investigate the circumstances surrounding the accident. The claims manager for Badger State, John Graeber, concluded after reading the police report and interviewing all of the automobile's occupants that the case was one of probable liability on McCarthy's part. He also determined that the case would probably involve damages to Mowry in excess of the \$16,000 policy limits.

Badger State's investigations indicated to it, however, that a question of policy coverage existed. The question revolved around the ownership of the vehicle which McCarthy was driving at the time of the accident. The insurer believed that it was unclear whether McCarthy or his parents were the true owners of the automobile involved in the accident. Its investigation disclosed that the car was titled in McCarthy's mother's name, but that McCarthy had paid for the car with his own money, did not need permission to drive the car, and had told several people at the scene of the accident that he owned the vehicle and that it was uninsured. Given these circumstances, Graeber concluded that the issue of ownership was debatable and that a serious question of coverage had arisen.

In March, 1976, ten months after the accident, Mowry filed suit against McCarthy, McCarthy's parents, Badger State, and an insurance agent. . . .

On September 13, 1977, Mowry issued a formal demand of settlement for the full amount of the liability insurance coverage, \$15,000. Badger State's attorney, Kurt Frauen, responded that Badger State had denied coverage under the policy, but that he would inform Badger State of the offer. Badger State did not accept the offer.

At a pretrial conference on September 26, 1977, Attorney Frauen requested that Mowry's counsel and other parties present agree to a bifurcated trial in which a determination on coverage would precede any trial on the issue of liability. In relating the events of the pretrial conference to Badger State, Attorney Frauen wrote, "Everyone seemed to feel that if the coverage issue were resolved, the rest of the case would not have to be tried." Mowry's counsel reiterated Mowry's offer of settlement on September 26, 1977, but Badger State again refused to accept it.

A subsequent stipulation and order set April 4, 1979, as the commencement date for the trial on the issue of coverage; the issues of liability and damages were to be held in abeyance until the resolution of the coverage issue. On March 12, 1979, approximately three weeks before the coverage trial, Mowry once more demanded that Badger State pay the limits of McCarthy's liability insurance policy plus \$1,000 under the medical payments coverage; he set a March 23, 1979, deadline for its acceptance. Badger State's counsel, in correspondence to Mowry's counsel, stated that he felt the settlement demand was contrary to the stipulation and order separating the coverage issue from the liability and damages issues: "The court has in fact bifurcated the trial . . . to resolve the coverage dispute before proceeding with the plaintiff's case." Mowry's counsel responded that he believed that stipulating to a bifurcation of issues "should not in any way be construed as barring plaintiff from attempting to negotiate settlement of his entire claim."

Badger State refused Mowry's March 12 settlement offer. On April 4, 1979, the coverage issue was tried before a jury. The jury returned a verdict the next day which found that McCarthy's parents owned the vehicle in question at the time of the accident. Coverage was thereby afforded Steven McCarthy under the policy.

On April 6, 1979, Badger State offered the limits on its liability policy and medical payments coverage. On January 10, 1980, Badger State's counsel informed McCarthy that it would assume McCarthy's defense in the action. Following negotiations between counsel for Mowry and Badger State, the parties entered into a stipulation of judgment in October, 1980, thereby rendering a trial on the liability and damages issues unnecessary. The judgment was in favor of Mowry and against Badger State for \$16,000 and against Steven McCarthy for \$175,000. The stipulation further called for McCarthy to assign to Mowry any and all causes of action which McCarthy might have against Badger State, in satisfaction of Mowry's judgment against McCarthy. Following that stipulation and entry of judgment, Mowry, suing under McCarthy's assignment of rights, brought the present action against Badger State for bad faith and breach of contract.

The circuit court, in holding that Badger State breached its contract in refusing to defend and that it committed bad faith in refusing to negotiate a settlement, was indignant that an insurer would delay settlement negotiations until the coverage issue has been judicially determined, particularly when liability and excess damages are undisputed. The court felt that Badger State's posture of not negotiating a settlement until the determination of the coverage issue placed all the risk of an excess judgment on the insured. It found that an insurance company who refuses to defend and refuses to negotiate may not protect itself from a claim for damages in excess of policy limits by tendering the policy limits only upon losing the coverage issue of a bifurcated trial. . . . Judge McGraw stated that the proper rule should be "when an offer of settlement within the policy limits has been made and ignored, a good faith refusal to defend is not a valid defense to a claim in excess of the policy limits. . . ." The court then awarded Mowry damages in the amount of \$159,000, representing the stipulated amount of liquidated damages for which Badger State would be liable in any action brought by Mowry as assignee against Badger State.

Badger State appealed. The court of appeals, in its certification memorandum, framed the

issue to be whether an insurance carrier "should be held liable for damages in excess of its policy coverage where its belief that there was no coverage led it to reject" an earlier offer of settlement within the policy limits. The court noted that this particular scenario presents an unaddressed area of insurance law in this state.

\* \* \*

We note the competing interests on each side of this case. When an insurer is certain of its insured's liability for an accident and where damages to the victim exceed policy limits, the insurer would normally be responsible for indemnifying its insured to the extent of its policy limits. The insurer, however, experiences a conflict of interests whenever an offer of settlement within policy limits is received where a legitimate question of coverage under the policy also exists. The insurer will be reluctant to settle within policy limits if there is a likelihood that coverage does not exist. On the other hand, an insurer's failure to settle a victim's claim within policy limits may subject an insured to a judgment in excess of his policy limits. This case presents a good example of these conflicting interests.

#### BAD FAITH CLAIM

An insurer owes a general duty to its insured to settle or compromise a claim made against the insured. This duty does not arise out of an express contractual provision; rather, it is implied from the terms of the contract which give the insurer the absolute control of the defense of the action against the insured.

The insurer has the right to exercise its own judgment in determining whether a claim should be settled or contested. But "exercise of this right should be accompanied by considerations of good faith." In order to be made in good faith, a decision not to settle a claim must be based on a thorough evaluation of the underlying circumstances of the claim and on informed interaction with the insured. This gives rise to several obligations on the part of the insurer. First, the insurer must exercise reasonable diligence in ascertaining facts upon which a good-faith decision to settle or not settle must be based. Second, where a likelihood of liability in excess of policy limits exists, the insurer must so inform the insured so that the insured might properly protect himself. Third, the insurer must keep the insured timely abreast of any settlement offers received from the victim and of the progress of settlement negotiations. . . . Mowry argues that the insurer's mistaken decision about the nonexistence of coverage should render Badger State liable for the excess judgment entered against the insured. He does not assert that Badger State breached any of the three traditional obligations arising out of the general duty to settle. Rather, he asserts that Badger State acted in bad faith in deciding to disclaim coverage where it was convinced that no real issue as to liability or damages existed. Moreover, he argues that Badger State's liability for damages caused by its refusal to settle an excess liability claim is not excused by a good-faith failure to defend.

\* \* \*

Mowry also cites *Comunale v. Traders & General Ins. Co.*, 50 Cal. 2d 654, 328 P.2d 198 (1958). . . .

The *Comunale rule*, in effect, renders an insurer strictly liable for any decision not to settle within policy limits, whether or not made in good faith, when a subsequent judgment against the insured exceeds policy limits. . . .

Although we acknowledge the apparent goal of the California approach — to protect the insured from liability for an excess judgment by placing the risk of an erroneous decision not to settle on an insurer — we decline to accept that strict approach for this jurisdiction. Such a policy is unduly oppressive on insurance companies and would force them to settle claims where

coverage may be dubious. The California approach is particularly unseemly in a jurisdiction such as our own, where an insurer may seek judicial determination of coverage issues prior to litigating liability and damages issues. The California approach is unrealistic if only to the extent that an insurer's belief that an insurance policy does or does not provide coverage must necessarily "affect a decision as to whether the settlement offer in question is a reasonable one."

\* \* \*

We hold that the circuit court erred in finding in this case that Badger State committed bad faith by refusing to settle and negotiate a settlement within policy limits. A finding of bad faith must not be measured solely against a backdrop that coverage was ultimately found to exist under the policy. Bad faith should be found in this case only if there was no fairly debatable coverage question. However, the circuit court, sitting without a jury, did not use this standard to reach its bad faith conclusion. The court concluded that Badger State's posture of waiting to defend the action and to negotiate a settlement until the coverage issue was determined itself constituted bad faith. Because the circuit court did not rely on appropriate and applicable law in making its bad faith determination, its holding is an abuse of discretion and, as such, is erroneous.

The upshot of the trial court's holding would be to require the insurer to accept an offer within policy limits even where a fairly debatable coverage question exists. We have, however, rejected the California approach, which would make an insurer strictly liable for an offer of settlement within policy limits.

Although this court might otherwise remand a matter to the circuit court for further consideration where the appropriate law has not been applied, we choose to decide the bad faith issue as a matter of law.

\* \* \*

We hold that Badger State did not act in bad faith in initially denying coverage to McCarthy. The record establishes a reasonable basis for Badger State's denial of coverage. Although the vehicle was registered in Mildred McCarthy's name, other items within the record suggest Steven McCarthy's ownership of the vehicle. Eugene McCarthy, Steven's father, stated that Steven paid for the car with money from Steven's own checking account, but that Eugene also "gave him some money." McCarthy apparently had approached his father about the idea of "buy[ing] the car" from a third party. Steven, according to the elder McCarthy, paid \$150 for the car, but Mildred McCarthy's name appeared on the title. McCarthy never had to ask his father or mother for permission to use the car. Eugene McCarthy suggested that his son obtain insurance on the vehicle. In a statement made to a Badger State representative, McCarthy claimed that his mother owned the car on the date of the accident but that he had paid for it. One of the vehicle's occupants on the date of the accident stated that McCarthy said at the accident scene that he owned the car and that it was uninsured.

Although the record itself does not conclusively establish that the vehicle in question was covered under the elder McCarthy's insurance policy (that task was undertaken and resolved by the jury in the coverage trial), the record sufficiently establishes that Badger State was presented with a fairly debatable coverage issue. There is no absence of a reasonable basis for Badger State's denial of coverage. The record gives no indication that Badger State failed to properly investigate the claim, or that important facts were recklessly ignored and disregarded. Badger State did not commit bad faith in failing to settle Mowry's claim within policy limits even though McCarthy's liability for the accident was probable and damages were concededly in excess of policy limits. The question of coverage under the policy was fairly debatable.

Badger State could have protected both its interests and its insured's interests by settling under a reservation of rights agreement, Mowry and the circuit court assert. The record reflects that Badger State's counsel and its claims manager considered such an option, but ultimately decided to pursue the bifurcation procedure. A reservation of rights agreement would result in a settlement of the injured's claims, while preserving the insurer's right to litigate the coverage issue. The insured benefits from this procedure because it is protected from excess judgment. Badger State argues, however, that the reservation of rights procedure will rarely result in the insurer's recouping the payments it made to the victim from the insured where coverage is found not to exist under the policy, because the insured may be judgment-proof.

An insurer is always at liberty to seek a reservation of its rights in settling a claim. But Badger State's failure to seek such a reservation in this case does not, by itself, constitute bad faith. Badger State merely sought a statutory mechanism to bifurcate the coverage issue from the liability and damages issues. To require an insurer to enter into a reservation of rights agreement in addition to proceeding within a separation framework would run contrary to the bifurcation allowed by [statute]. Even though Badger State's determination that coverage did not exist was wrong, its mistake does not mean that it acted in bad faith in refusing to settle if the issue of coverage was fairly debatable. . . .

Mowry argues that it is inherently unfair for McCarthy to be liable for a judgment in excess of policy limits when the judgment could have been wholly avoided had Badger State settled within policy limits when it had the opportunity to do so. But to require an insurer to settle any claim within policy limits where the insured's liability and the victim's damages in excess of policy limits are relatively certain, without consideration as to whether coverage exists, may result in extortionate lawsuits against the insurer. . . . The bad faith standard . . . strikes an acceptable balance between the insurer's and insured's competing interests concerning settlement offers within policy limits where liability and damages in excess of policy limits are apparent.

\* \* \*

Because the trial court held that Badger State committed bad faith in refusing to negotiate or settle a claim within its policy limits and did not apply the fairly debatable standard to the coverage question, the trial court erred as a matter of law. We hold that when a coverage issue is fairly debatable, an insurer will not have acted in bad faith in refusing to settle within policy limits, even when the insured's liability for the incident is undisputed and when the victim's damages appear to exceed policy limits. Because no determination was made that the coverage issue was fairly debatable, the circuit court abused its discretion in holding that Badger State acted in bad faith and in holding it liable for McCarthy's excess judgment.

\* \* \*

### **Notes and Questions**

**1. Specifying the Conflict of Interest:** *Mowry* presents a clear conflict of interest between the insured and the insurer. The parties face a situation in which the insured's liability is near certain and there is a high probability that damages will be in excess of insurance coverage. Thus, the insured has a strong incentive to take any settlement offer that is within the coverage limits. However, to the extent that there is an issue regarding coverage, the insurer's incentives deviate from those of the insured. Specifically, if the insurer pays the settlement and later wins on the coverage issue, the chance of recovering the amount paid in settlement from the insured is remote. Therefore, the insurer prefers to decide the coverage issue first. In short, the issue in *Mowry* is who bears the risk that, by refusing settlement, damages in excess of the policy limits will be awarded. Do the concepts of freedom of contract and Coasian bargaining have any

implications for your analysis of this conflict of interest?

**2. *Strict Liability Duty to Settle:*** Under a strict liability rule, the insurer always bears the risk of an excess judgment. Insurers could exercise reservation rights, but the value of such rights is diminished by the low probability of ex post collection. The court observes that such a rule provides an incentive for "extortionate" lawsuits. Why does the court state that the rule provides that incentive? How is this so? Moreover, notice that a strict liability rule constrains the ability of an insurer to negotiate a settlement. In other words, the victim can offer to settle at the maximum required under the policy with a large degree of confidence that the offer will be accepted. Why is this so? How does a strict liability rule impact the cost of providing insurance? From a public policy perspective, why should cost matter?

**3. *Good Faith Duty to Settle:*** Under a good faith rule, the insurer can shift the risk of an excess judgment on to the insured. Does this prevent "extortionate" lawsuits and provide a greater degree of flexibility in negotiating settlements? Does a good faith rule place the risk of an excess judgment on the party who ex ante was the least cost avoider? How does a good faith rule impact the cost of providing insurance? From a public policy perspective is this a better result than under a strict liability rule?

### **3. Self-Insurance**

Rather than purchasing insurance, some individuals decide to bear the risk themselves. That is, they self insure. Self-insurance creates incentives to do a number of different things. First, individuals can invest resources to minimize the probability of an uncertain event occurring. Examples of this would be an individual clearing the brush away from the area surrounding their house to minimize the likelihood that their house will burn down in the event of a brush fire. Second, self-insurance also creates incentives to invest in security precautions to minimize the monetary loss in the event the contingency actually occurs. Thus, individuals could install a sprinkler system, fire alarms, and smoke detectors with telephone links to the local fire station to minimize the monetary loss in the event that their dwelling catches on fire. Third, self-insurance gives individuals the incentive to set aside reserve funds to cover possible losses. Many large companies self insure their employees' health insurance. The companies often hire a third-party administrator to administer the self-insurance process within the company in order to save on transactions costs and to take advantage of the administrator's expertise.

### **C. Risk and Market Prices**

Markets help individuals and firms deal with uncertainty and risk by adjusting and discounting the market price to reflect the risk. This is clearly seen in the case of financial products where riskier stocks receive a higher return than less risky stocks. Similarly, unsafe products or services are often sold at a discount. For example, many people (perhaps mistakenly) believe that some of the discount airlines are not as safe as the major airlines and, thus, they are not willing to fly on the discount airlines even though the price is lower. One way to characterize the market adjustment process here is that the low fare airlines must reduce fares to get people to take on the perceived increased risk associated with flying on those airplanes. In general, the market demand curve shifts back to the left as consumers react adversely to labeling of a product as "unsafe." The result is that unsafe products (that is, products individuals perceive to be riskier than other products) sell at a discounted price relative to the price of safer products.

### **1. Financial Products**



Financial products offer perhaps the clearest evidence of market adjustment to risk. Investors typically require higher expected returns as compensation for bearing greater risks. Consider two alternative investments, A and B. A and B have the same expected annual return for each of the next ten years and thus have the same net present value. However, A offers a risk-free rate of return, and B has considerable variance in the possible cash flow in each year. In order to induce investors to purchase asset B, it must be sold at a discounted price relative to A. By forcing the purchase price of B down, the return on that investment increases. Therefore, the market forces a higher expected return for the riskier investment.

### **The Rights of Creditors of Affiliated Corporations**

Richard A. Posner

43 U. Chi. L. Rev. 499, 501–03 (1976)

Mr. A. Smith wants to borrow \$1 million to invest in a mining venture together with \$2 million of his own money. He wants the loan for only a year since by the end of the year it will be apparent whether the venture has succeeded; if it has, he would then want to obtain longer-term financing. Since Smith is a man of means, if he gives his personal note to the lender the latter would regard a one-year loan of \$1 million as riskless and would offer Smith the riskless short-term interest rate, say six percent. But Smith is reluctant to stake more than \$2 million on the outcome of the mining venture. He proposes to the lender a different arrangement, whereby the lender will agree to look for repayment of the loan exclusively to the assets of the mining venture, if any exist, a year hence. Under this arrangement, Smith will be able to limit his liability to his investment in the venture.

The lender estimates that there is an 80 percent probability that the venture will be sufficiently successful to enable repayment of the loan and interest on the due date, and a 20 percent probability that the venture will fail so badly that there will be insufficient assets to repay even a part of the loan.<sup>1</sup> On these assumptions the solution to the lender's problem is purely mechanical; he must calculate the amount, payable at the end of a year, that when multiplied by 80 percent (the probability that payment will in fact be made) will equal \$1,060,000, the repayment he would have received at the end of the year had he made the riskless loan. That amount is \$1,325,000.<sup>2</sup> Accordingly, the lender will charge Smith 32.5 percent interest for the loan if Smith's obligation to repay is limited to the assets of the venture. At this rate of interest the lender is indifferent as between the riskless and the risky loan.

This example illustrates the fundamental point that the interest rate on a loan is payment not only for renting capital but also for the risk that the borrower will fail to return it. It may be wondered why the borrower might want to shift a part of the risk of business failure to the lender, given that he must compensate him for bearing added risk. There are two reasons why the lender might be the superior risk bearer. First, the lender may be in a better position than the borrower to appraise the risk. Compare the positions of the individual shareholder in a publicly held corporation and the banks that lend the corporation its working capital. It may be easier and hence cheaper for the bank to appraise the risk of a default and the resulting liability than it would be for the shareholder, who may know little or nothing about the business in which he has invested. Second, the borrower may be risk averse and the lender less so (or risk neutral, or even

---

<sup>1</sup> Plus the additional assumption that the lender is "risk neutral." This term is explained in note [3] infra.

<sup>2</sup> Calculated by solving  $.8x = \$1,060,000$  for  $x$ .

risk preferring).<sup>3</sup> Thus, unlimited liability would discourage investment in business ventures by individuals who wanted to make small, passive investments in such ventures. It would also discourage even substantial entrepreneurial investments by risk-averse individuals — and most individuals are risk averse.

A borrower could in principle negotiate with the lender for an express limited-liability provision. The more usual course, however, is to incorporate and have the corporation borrow the money. The basic principle of corporation law is that the shareholders of a corporation are not personally liable for the corporation's debts unless they agree to assume such liability. Corporate borrowing therefore automatically limits the borrower's liability to his investment in the corporation. The fact that the law permits Smith to limit his liability by conducting his mining venture in the corporate form does not imply, however, that the law is somehow tilted against creditors or enables venturers to externalize the risks of business failure. . . . Although incorporation permits Smith to shift a part of the risk of failure to the lender, there is no externality; the lender is fully compensated by the higher interest rate that the corporation must pay by virtue of enjoying limited liability. Moreover, the lender is free to insist as a condition of making the loan that Smith guarantee the debts of the corporation personally or that he consent to including in the loan agreement other provisions that will limit the lender's risk — though any reduction in the risk will reduce the interest rate the lender can charge since a portion of that rate is, as we have seen, compensation to the lender for agreeing to bear a part of the risk of the venture.

There is an instructive parallel here to a fundamental principle of bankruptcy law: the discharge of the bankrupt from his debts. This principle, which was originally developed for the protection of business rather than individual bankrupts, enables the venturer to limit his risk of loss to his current assets; he is not forced to hazard his entire earning capacity on the venture. Incorporation performs the same function of encouraging investment by enabling the risk averse to limit their risk of loss to their investment.

Far from externalizing the risks of business ventures, the principle of limited liability in corporation law facilitates a form of transaction advantageous to both investors and creditors; in its absence the supply of investment and the demand for credit might be much smaller than they are. . . .

## **2. Products and Services: Risk Allocation and Contract Law**

Prices for products and services adjust to reflect the risk associated with the transaction. In essence, freedom of contract allows parties to allocate risk according to which party is willing

---

<sup>3</sup>.An individual is risk averse if he prefers the certain equivalent of an expected value to the expectation — the certainty of receiving \$1 to a 10 percent chance of receiving \$10. A risk preferrer would have the opposite preference, and a risk-neutral individual would be indifferent. A corporation is less likely to be risk averse than an individual because the shareholders of the corporation can offset any risks incurred by that corporation by holding a diversified portfolio of securities. Moreover, a large lender can eliminate or greatly reduce the risk of loss on a particular loan by holding a diversified portfolio of loans. On both counts it seems likely that individual investors would often be more averse to bearing unlimited personal liability for the failure of an enterprise in which they had invested than lenders would be to bearing the risk of that failure to the extent of their loan to the enterprise.

to accept the risk in light of the price adjustment. This subsection considers various areas where contracts allocate risk.

*a. Assumption of Risk*

Individuals often engage in activities that involve the risk of injury. As a defense to a negligence claim, the doctrine of assumption of the risk prohibits recovery for injuries that an individual receives when he voluntarily exposes himself to a known and appreciated danger. Four elements are required to establish an assumption of the risk defense. The plaintiff must: (1) have knowledge of facts which constitute a dangerous condition, (2) know that the condition is dangerous, (3) appreciate the nature or extent of the danger, and (4) voluntarily expose himself to the danger. Although in some cases, a written contract expressly provides for an assumption of the risk, the aforementioned elements make it clear that an implied assumption of the risk is comparable to an implied contractual agreement.

Suppose, for example, that two electric drills available in a hardware store are priced at \$20 and \$50. It is explicitly stated on the package of the lower priced drill that it is not safe for use in drilling through concrete. A purchaser of the lower priced drill who plans to drill through concrete does so because the added protection of the more expensive drill is not worth the added cost. Thus, the purchaser assumes the risk of injury that may occur when drilling through concrete. Under contract law and tort law, the purchaser may not be able to recover damages from the manufacturer because voluntarily bearing the risk of injury often serves as a bar to recovery. However, a person can only assume known and appreciated risks. An example often given is a spectator at a baseball game assumes a certain amount of risk that he or she will be injured if hit by a baseball. A spectator probably assumes the risk of being hit by a baseball in areas that are not protected by screens because most can appreciate the risk they are taking. However, spectators do not assume the risk that a ball will penetrate a protective screen and cause them harm since it is difficult to assess the risk of such an event.

**Avram Wisna v. New York University**  
Supreme Court of New York  
Docket No. 114439/2005 (Jan. 23, 2008)

CAROL R. EDMEAD, Judge.

\* \* \*

The following facts are not in dispute. During the spring of 2004, plaintiff Avram Wisnia was a matriculated student at NYU, completing his junior year. He resided at an NYU residence hall, Third Avenue North, that is located at 75 Third Avenue, New York, New York (the Premises). During his junior year, plaintiff was elected as a Secretary of the Third Avenue North Student Counsel (TASC), a student-run organization responsible for, among other things, planning and overseeing dormitory events. As part of its spring activities, TASC organized a "Beach Bash Event" to take place in the courtyard of the Premises on May 1, 2004. The activities on that day included: a DJ, a "moon-bounce," jell-o wrestling in a kiddie pool, volleyball, water guns, and water balloons. Plaintiff took part in planning the Beach Bash. He was responsible for advertising the event to other students living in the residence hall and he was in charge of supplying food for the event.

On May 1, 2004, plaintiff alleges that he took a trip to the store to purchase snacks for the day's events. Upon his return to the Premises, several of his friends from TASC beckoned him to

investigate the quality of the jell-o in the kiddie pool. Plaintiff placed the items he purchased from the store on a nearby table and proceeded to walk over to look at the pool. When he arrived at the pool, two members of TASC, Alex and Carmen, grabbed plaintiff and pushed him into the kiddie pool. Plaintiff climbed out of the pool, removed his cell phone and wallet from his pocket, walked over to Carmen and started grappling with the young man. At that time, they were in close proximity to the kiddie pool. After several seconds of horse-play, both men landed in the kiddie pool after being pushed in by another member of TASC. It is during this second fall that plaintiff allegedly sustained an injury to his hip.

The entire incident was captured on video by a DJ hired to video record the event. As a result, plaintiff commenced this lawsuit asserting a cause of action for negligence, asserting damages in the amount of one million dollars.

\* \* \*

### ***Assumption of the risk***

Despite plaintiff's argument to the contrary, it is difficult to imagine a more compelling set of facts for the application of the doctrine of primary assumption of risk. That doctrine provides that a voluntary participant in a sporting or recreational activity "consents to those commonly appreciated risks which are inherent in and arise out of the nature of the sport [or recreational activity] generally and flow from such participation." Contrary to defendant's contention, individuals are not restricted to playing organized sports for this doctrine to apply. The doctrine also applies to various types of unorganized sports as well as many forms of recreational activity.

A plaintiff who voluntarily participates in a recreational activity is deemed to consent to the apparent or reasonably foreseeable consequences of that activity. This includes those risks associated with the construction of the playing surface and any open and obvious condition on it. Moreover, it is not necessary to the application of the doctrine that the injured plaintiff may have foreseen the exact manner in which the injury occurred, so long as he or she is aware of the potential for injury from the mechanism from which the injury results. Therefore, "if the risks of the activity are fully comprehended or perfectly obvious, the plaintiff has consented to them and the defendant has performed its duty."

In his examination before trial, plaintiff testified that he took part in planning the events of the Beach Bash, which included a DJ, moon-bounce, jell-o wrestling in a kiddie pool, waterguns, water-balloons, and volleyball. In said testimony, plaintiff acknowledged that he was aware of the existence of, and the dangerous nature of, conducting a jell-o wrestling match in a kiddie pool on the concrete surface of the courtyard. In fact, it was plaintiff who suggested placing gym mats underneath the pool to provide added safety.

However, plaintiff argues that he did not voluntarily enter into the kiddie pool, and thus should not be deemed to have assumed the risk for his injuries. This Court is not persuaded by plaintiff's argument. Plaintiff voluntarily entered the Premises to take part in the activities that were planned on that day. Those activities included a jell-o-wrestling match. The record also reveals that, after the first roughhousing incident that landed plaintiff into the kiddie pool, plaintiff climbed out of the pool, took his cell phone and wallet out of his pocket, placed them on a table, approached Carmen (who was standing in front of the kiddie pool), grappled with him and they were subsequently pushed into the pool a second time. Plaintiff testified that it was this second incident that caused his injuries. Thus, it is clear from plaintiff's testimony that he voluntarily engaged in grappling, wrestling, and rough play on the concrete surface in front of the pool. Although plaintiff may not have foreseen the exact manner in which his injury

occurred, he was aware of the potential for injury when he decided to grapple with another student in front of a kiddie pool on a concrete surface. Under these facts, plaintiff fully comprehended the risk of the activity in which he was engaged and assumed the risk of the injuries which he sustained.

### **Notes and Questions**

**1. Risk Preferences and Assumption of Risk:** In general, it is costly for a defendant to satisfy a duty of care owed to a plaintiff. Often, the costs of providing the level of care indicated by tort are passed on to consumers in the form of higher prices, entrance fees, rates, etc. However, some consumers might be willing to release the potential tortfeasor from its duty in order to obtain a lower price. The assumption of risk doctrine allows parties to contract to bear a risk when they have a comparative advantage in risk bearing. That is, the assumption of risk doctrine facilitates voluntary, mutually beneficial exchange.

**2. Least Cost Avoiders and Assumption of Risk:** Economic efficiency is enhanced if risk is reduced in the least costly manner. In the absence of the assumption of risk doctrine, defendants bear the cost of reducing risk regardless of whether they have a comparative advantage. In some instances, the party to whom a duty is owed may be able to insure against the occurrence of particular risks more cheaply than the defendant. Such parties would find it in their economic self-interest to relieve the defendant of its duty of care by express agreement.

#### ***b. The Bargaining Principle and Least-Cost Risk Avoider***

As discussed in Chapter V, contract law reduces transaction costs when it supplies default terms for contracts that are the terms for which most contracting parties would bargain if negotiation were costless. This **bargaining principle** provides a framework for analyzing numerous contract law rules, as well as rules in other substantive areas of law such as bankruptcy and corporate law. Although it is not clear what terms most bargaining parties would negotiate for in the absence of negotiation costs, economics can derive some general principles from assumptions about behavior and mutually beneficial exchange. Most economic actors are risk averse — they are willing to pay (or accept a lower price) in order to increase certainty. Moreover, because negotiating parties wish to maximize the total gain from the exchange, they will allocate risk to the lowest-cost risk avoider (i.e., the party with the comparative advantage at avoiding or bearing risk) if transaction costs are zero. This analysis has implications for how courts should interpret contracts when an unforeseeable contingency occurs — courts should allocate risk according to how the contracting parties would have allocated the risk themselves if they had known of the risk at the time of contracting. This determination requires analysis of the relative risk-bearing abilities of the contracting parties. In general, the party with the comparative advantage at bearing a particular risk — the so-called least-cost risk avoider — will be required to bear the risk.

#### ***c. The Allocation of Risk — Impracticability and Mistake***

The ex ante expected mutual benefits of a contract may change as a result of the occurrence of some unforeseen contingency, sometimes causing the contract to appear unfair ex post. Commercial impracticability and mistake (mutual or unilateral) are the defenses under which a court might excuse a promisor from performance under such changed circumstances. Impracticability and mistake cases can be distinguished as follows: Impracticability deals with a performance made burdensome by an exogenous event (such as a dramatic, unprecedented

increase in prices caused by some unpredictable event beyond the control of the contracting parties). Mistake, on the other hand, deals with the materialization of an endogenous risk (such as the parties' realization that the subject of the contract is of a different make or model of automobile than they had thought). As with any potential excuse from performance, courts must be careful to make sure an opportunistic promisor does not use the excuse to take advantage of the changed circumstances.

**Eastern Air Lines v. Gulf Oil Corp.**

United States District Court for the Southern District of Florida  
415 F. Supp. 429 (1975)

JAMES LAWRENCE KING, District Judge.

Eastern Air Lines, Inc., hereafter Eastern, and Gulf Oil Corporation, hereafter Gulf, have enjoyed a mutually advantageous business relationship involving the sale and purchase of aviation fuel for several decades.

This controversy involves the threatened disruption of that historic relationship and the attempt, by Eastern, to enforce the most recent contract between the parties. On March 8, 1974, the correspondence and telex communications between the corporate entities culminated in a demand by Gulf that Eastern must meet its demand for a price increase or Gulf would shut off Eastern's supply of jet fuel within fifteen days.

Eastern responded by filing its complaint with this court, alleging that Gulf had breached its contract and requesting preliminary and permanent mandatory injunctions requiring Gulf to perform the contract in accordance with its terms. By agreement of the parties, a preliminary injunction preserving the status quo was entered on March 20, 1974, requiring Gulf to perform its contract and directing Eastern to pay in accordance with the contract terms, pending final disposition of the case.

Gulf answered Eastern's complaint, alleging that the contract was not a binding requirements contract, was void for want of mutuality, and, furthermore, was "commercially impracticable" within the meaning of Uniform Commercial Code § 2-615.

[After determining that the requirements contract was enforceable and that Eastern had not breached it, the court addressed the § 2-615 issue.]

In short, for U.C.C. § 2-615 to apply there must be a failure of a pre-supposed condition, which was an underlying assumption of the contract, which failure was unforeseeable, and the risk of which was not specifically allocated to the complaining party. The burden of proving each element of claimed commercial impracticability is on the party claiming excuse.

The modern U.C.C. § 2-615 doctrine of commercial impracticability has its roots in the common law doctrine of frustration or impossibility and finds its most recognized illustrations in the so-called "Suez Cases," arising out of the various closings of the Suez Canal and the consequent increases in shipping costs around the Cape of Good Hope. Those cases offered little encouragement to those who would wield the sword of commercial impracticability. As a leading British case arising out of the 1957 Suez closure declared, the unforeseen cost increase that would excuse performance "must be more than merely onerous or expensive. It must be positively unjust to hold the parties bound." *Ocean Tramp Tankers v. V/O Sovfracht (The Eugenia)*, 2 Q.B. 226, 239 (1964). . . .

Other recent American cases similarly strictly construe the doctrine of commercial impracticability. For example, one case found no U.C.C. defense, even though costs had doubled over the contract price, the court stating, "It may have been unprofitable for defendant to have

supplied the pickers, but the evidence does not establish that it was impossible. A mere showing of unprofitability, without more, will not excuse the performance of a contract."

Recently, the Seventh Circuit has stated: "The fact that performance has become economically burdensome or unattractive is not sufficient for performance to be excused." We will not allow a party to a contract to escape a bad bargain merely because it is burdensome. "[The] buyer has a right to rely on the party to the contract to supply him with goods regardless of what happens to the market price. That is the purpose for which such contracts are made. . . ."

...

[The court held that Gulf had not shown the hardship required for performance to be commercially impracticable.]

But even if Gulf had established great hardship under U.C.C. § 2-615, which it has not, Gulf would not prevail because the events associated with the so-called energy crises were reasonably foreseeable at the time the contract was executed. If a contingency is foreseeable, it and its consequences are taken outside the scope of U.C.C. § 2-615, because the party disadvantaged by fruition of the contingency might have protected himself in his contract.

\* \* \*

The record is replete with evidence as to the volatility of the Middle East situation, the arbitrary power of host governments to control the foreign oil market, and repeated interruptions and interference with the normal commercial trade in crude oil. Even without the extensive evidence present in the record, the court would be justified in taking judicial notice of the fact that oil has been used as a political weapon with increasing success by the oil-producing nations for many years, and Gulf was well aware of and assumed the risk that the OPEC nations would do exactly what they have done.

#### IV. REMEDY

Having found and concluded that the contract is a valid one, should be enforced, and that no defenses have been established against it, there remains for consideration the proper remedy. . . . Gulf presently supplies Eastern with 100,000,000 gallons of fuel annually or 10 percent of Eastern's total requirements. If Gulf ceases to supply this fuel, the result will be chaos and irreparable damage. . . . It has previously been found and concluded that Eastern is entitled to Gulf's fuel at the prices agreed upon in the contract. In the circumstances, a decree of specific performance becomes the ordinary and natural relief rather than the extraordinary one. The parties are before the court, the issues are squarely framed, they have been clearly resolved in Eastern's favor, and it would be a vain, useless and potentially harmful exercise to declare that Eastern has a valid contract, but leave the parties to their own devices. Accordingly, the preliminary injunction heretofore entered is made a permanent injunction and the order of this court herein. \* \* \*

### Notes and Questions

**1. *Ex Ante and Ex Post Fairness:*** Reciprocal benefit is the basis of contractual interaction. Each party agrees to the contract terms in anticipation of benefiting from the other's performance. The fairness of contract terms can be evaluated from two perspectives. An *ex ante* perspective views the terms at the time of agreement, *before* unknown or unanticipated contingencies occur. *Ex ante* fairness reflects the mutually beneficial aspect of contractual exchange — if both parties agreed to and understood the terms, then the contract is fair. *Ex post* fairness means evaluating the contract after the agreement and after the occurrence of contingencies. Some of the most interesting problems in contract law deal with the question of

when courts should not enforce contract terms agreed to ex ante because of the occurrence of events ex post. Recognition of the important roles of contract law in enforcing contract terms and allocating risk ex ante reveals the potential problems of ex post judicial reconstruction of contracts. Moreover, the fairness or unfairness of particular contract terms should not be judged in isolation, but should be evaluated in light of all aspects of the bargaining process.

**2. *Least-Cost Risk Avoider:*** As a starting point for the analysis of the efficiency of contractual risk allocation, it must be recognized that some parties in the economy possess a comparative advantage at minimizing the costs associated with risk. Such parties are referred to as the "least-cost avoider of risk." Good examples of such people are grain dealers, who buy and sell contracts for the future delivery of grain. In essence, grain dealers minimize risk by engaging in numerous buy and sell contracts so an unexpected fluctuation — either up or down — in the market price of grain is not disastrous to them. Moreover, the grain dealers specialize in understanding the grain market and are able to hedge their risks on the basis of their specialized information. On the other hand, the individual farmer does not have the expertise to engage in the pooling of risks on his or her own, but may lock into a guaranteed price long before the market price at harvest is determined. The farmer clearly benefits from not exposing his or her entire income stream to a last minute fluctuation in the market price of grain. Put another way, the grain dealer specializes in maintaining a diversified portfolio while the individual farmer has a non-diversified portfolio. The grain dealer sells insurance (i.e., lower risk) to the risk-averse farmer. Understanding and appreciation of the risk allocation role of contracts leads to richer analyses of various doctrines of contract law. In general, legal rules that enforce contracts when the risk is borne by the least-cost avoider of risk are efficient.

**3. *Cost Minimization and Least-Cost Risk Avoider:*** At the time of contracting, the parties have a mutual interest in reducing the costs associated with the contract as much as possible. If the parties consider the potential risks that might thwart performance or make performance impractical, they would place the risk on the party best able to deal with it — that is, the party with the comparative advantage in buying insurance or foreseeing or taking appropriate precautions to reduce the costs of the risk. In general, in the absence of a specific agreement, contract law imposes the risk on the party that seems to have the comparative advantage in bearing risk. Indeed, this is how similarly situated parties would be expected to allocate risks if required to bargain explicitly in advance. By assigning risks of nonperformance to the party who can best control the performance, contract law encourages parties to consider risk when negotiating a contract.

**Wilkin v. 1st Source Bank**  
Court of Appeals of Indiana  
548 N.E.2d 170 (1990)

HOFFMAN, J.

Respondents-appellants Terrence G. Wilkin and Antoinette H. Wilkin (the Wilkins) appeal from the judgment of the St. Joseph Probate Court in favor of petitioner-appellee 1st Source Bank (Bank). The Bank, as personal representative of the estate of Olga Mestrovic, had filed a petition to determine title to eight drawings and a plaster sculpture owned by Olga Mestrovic at the time of her death but in the possession of the Wilkins at the time the petition was filed. The probate court determined that the drawings and sculpture were the property of the estate, and the court ordered the Wilkins to return the items to the Bank.



At the request of the Bank, the probate court entered findings of fact and conclusions of law. Neither party disputes the validity of the findings of fact. Accordingly, this Court will accept the findings as true. The findings of fact may be summarized as follows.

Olga Mestrovic died on August 31, 1984. Her last will and testament was admitted to probate on September 6, 1984, and the Bank was appointed personal representative of the estate.

At the time of her death, Olga Mestrovic was the owner of a large number of works of art created by her husband, Ivan Mestrovic, an internationally-known sculptor and artist.<sup>4</sup> By the terms of Olga's will, all the works of art created by her husband and not specifically devised were to be sold and the proceeds distributed to members of the Mestrovic family.

Also included in the estate of Olga Mestrovic was certain real property. In March of 1985, the Bank entered into an agreement to sell the real estate to the Wilkins. The agreement of purchase and sale made no mention of any works of art, although it did provide for the sale of such personal property as the stove, refrigerator, dishwasher, drapes, curtains, sconces and French doors in the attic.

Immediately after closing on the real estate, the Wilkins complained that the premises were left in a cluttered condition and would require substantial cleaning effort. The Bank, through its trust officer, proposed two options: the Bank could retain a rubbish removal service to clean the property or the Wilkins could clean the premises and keep any items of personal property they wanted. The Wilkins opted to clean the property themselves. At the time arrangements were made concerning the cluttered condition of the real property, neither the Bank nor the Wilkins suspected that any works of art remained on the premises.

During their clean-up efforts, the Wilkins found eight drawings apparently created by Ivan Mestrovic. They also found a plaster sculpture of the figure of Christ with three small children. The Wilkins claimed ownership of the works of art, based upon their agreement with the Bank that if they cleaned the real property then they could keep such personal property as they desired.

The probate court ruled that there was no agreement for the purchase, sale or other disposition of the eight drawings and plaster sculpture. According to the lower court, there was no meeting of the minds, because neither party knew of the existence of the works of art.

On appeal, the Wilkins contend that the court's conclusions of law were erroneous. . . .

Mutual assent is a prerequisite to the creation of a contract. Where both parties share a common assumption about a vital fact upon which they based their bargain, and that assumption is false, the transaction may be avoided if because of the mistake a quite different exchange of values occurs from the exchange of values contemplated by the parties. There is no contract, because the minds of the parties have in fact never met.

---

<sup>4</sup> Ivan Mestrovic lived from 1883 to 1962. A Yugoslavian sculptor, Mestrovic became internationally known and spent his final working years in the United States. He taught at Syracuse University and at the University of Notre Dame. Much of his instruction as well as his own creative work involved religious themes.

Mestrovic is represented in the United States in the Brooklyn Museum, Art institute of Chicago, Colgate University, Syracuse University, Syracuse Museum of Fine Arts and University of Notre Dame; in England, in the Leeds Art Gallery, Birmingham Museum of Art, Victoria and Albert Museum and Tate Gallery; and in various collections in Canada, Spain, Belgium and Italy. A great number of churches in various parts of the world own his works.

The necessity of mutual assent, or "meeting of the minds," is illustrated in the classic case of *Sherwood v. Walker* (1887), 66 Mich. 568, 33 N.W. 919. The owners of a blooded cow indicated to the purchaser that the cow was barren. The purchaser also appeared to believe that the cow was barren. Consequently, a bargain was made to sell at a price per pound at which the cow would have brought approximately \$80.00. Before delivery, it was discovered that the cow was with calf and that she was, therefore, worth from \$750.00 to \$1,000.00. The court ruled that the transaction was voidable:

"The mistake was not of the mere quality of the animal, but went to the very nature of the thing. A barren cow is substantially a different creature than a breeding one. There is as much difference between them . . . as there is between an ox and a cow. . . ."

Like the parties in *Sherwood*, the parties in the instant case shared a common presupposition as to the existence of certain facts which proved false. The Bank and the Wilkins considered the real estate which the Wilkins had purchased to be cluttered with items of personal property variously characterized as "junk," "stuff" or "trash." Neither party suspected that works of art created by Ivan Mestrovic remained on the premises.

As in *Sherwood*, one party experienced an unexpected, unbargained-for gain while the other party experienced an unexpected, unbargained-for loss. Because the Bank and the Wilkins did not know that the eight drawings and the plaster sculpture were included in the items of personalty that cluttered the real property, the discovery of those works of art by the Wilkins was unexpected. The resultant gain to the Wilkins and loss to the Bank were not contemplated by the parties when the Bank agreed that the Wilkins could clean the premises and keep such personal property as they wished.

The following commentary on *Sherwood* is equally applicable to the case at bar:

"Here the buyer sought to retain a gain that was produced, not by a subsequent change in circumstances, nor by the favorable resolution of known uncertainties when the contract was made, but by the presence of facts quite different from those on which the parties based their bargain."

Palmer, *Mistake and Unjust Enrichment* 16–17 (1962). The probate court properly concluded that there was no agreement for the purchase, sale or other disposition of the eight drawings and plaster sculpture, because there was no meeting of the minds.

The judgment of the St. Joseph Probate Court is affirmed.

### **Notes and Questions**

**1. *Mistake or Ex Ante Risk Allocation?*:** The defense of mistake applies to situations where there was a genuine misunderstanding of the facts at the time the contract was made, not where some contingency occurred after the contract was made. Therefore, the fundamental issue is not risk allocation, but rather whether the parties actually agreed. Subsequent legal research on the facts of *Sherwood v. Walker* has revealed that, at the time of contract, both parties recognized that there was some probability that the cow, Rose 2d of Aberlone, was not barren. If this is correct, then didn't the contract terms reflect the probability that Rose 2d could become pregnant? That is, didn't the parties agree to shift the risk? If so, then why should the contract be set aside and the seller allowed to keep the gain that resulted from the occurrence of the highly unlikely contingency? Moreover, since the seller was in a better position to calculate the probability of the contingency occurring, doesn't the decision in the case represent an unwarranted windfall for him?

**2. Mistake, Fraud, Disclosure, and Regulation:** The producer or seller typically has an informational advantage over the buyer. A market-oriented response to concerns that sellers will use their informational advantage to abuse purchasers is that the market will eventually punish a seller for such indiscretions. See the discussion of reputational sanctions in Chapter V. While generally effective, the process of market adjustment is inadequate to address the harm caused by the first occurrence of fraud, deception, or accident due to unsafe design. For example, an individual who has been permanently maimed by an unsafe product takes little pleasure in knowing that the market will adjust to prevent the future sale of the company's unsafe products. Thus, the market adjustment process is supplemented by common law rules and statutory regulations. Indeed, government intervention is often justified when informational asymmetries are such that consumers, investors, and workers are exposed to risks of which they are unaware. Such government intervention can take the form of disclosure regulation (e.g., federal securities regulation) or direct regulation of acceptable risks (e.g., the Occupational Safety and Health Administration's workplace regulations, or the Food and Drug Administration's drug approval process). Market adjustments to different types of risk are discussed in the following section.

### **3. Compensating Wage Differentials and Market Levels of Safety**

The theory of compensating wage differentials can provide many insights into how the market place deals with risk and perceptions of risk. In order to isolate the impact of risk of injury on the job, it is necessary to assume that the compensating wage differentials for every other dimension of the job have already been established. To obtain a complete understanding of the job selection process and its outcomes, it is necessary, as always, to consider both the employer and the employee sides of a market.

Life is full of tradeoffs. Workers are often willing to trade the risk of injury for the receipt of higher wages. The worker who is offered a job for \$10.00 per hour in a business firm in which four percent of the work force is injured each year would achieve a certain level of utility from that job. If the risk of injury increased to five percent holding the other job characteristics constant, then the job would have to pay a higher wage to produce the same level of utility. In this regard, compensating wage differentials provide an ex ante compensation related to injury risk. People, of course, differ in their aversion to the risk of being injured. Those who are very sensitive to risk require large wage increases for any increase in risk, while those who are less sensitive require smaller wage increases to hold utility constant.

Workers can also be compensated to keep utility constant by ex post payments for damages. Workers' compensation insurance provides for ex post payments, but these payments are typically incomplete. There is no way to compensate a worker for his or her own death, and workers' compensation does not cover the psychic costs of disfigurement due to permanent impairment. Moreover, the lost income associated with temporary impairments is not fully replaced by workers' compensation because not all injury-related losses are completely compensated ex post. The advantage of ex post compensation comes from the difficulty of estimating some risks. The combination of ex ante compensating wage differentials and ex post damage payments holds worker utility constant in the face of increased risk.

Employers face a wage risk tradeoff of their own. In general, competitive forces on the employers' side of the market tend to cause low risk to be associated with low wages and high risk to be associated with higher wages, holding other things constant. This conclusion follows from three assumptions. First, it is presumably costly to reduce the risk of injury to employees. Safety equipment must be placed on machines, production time must be sacrificed for training

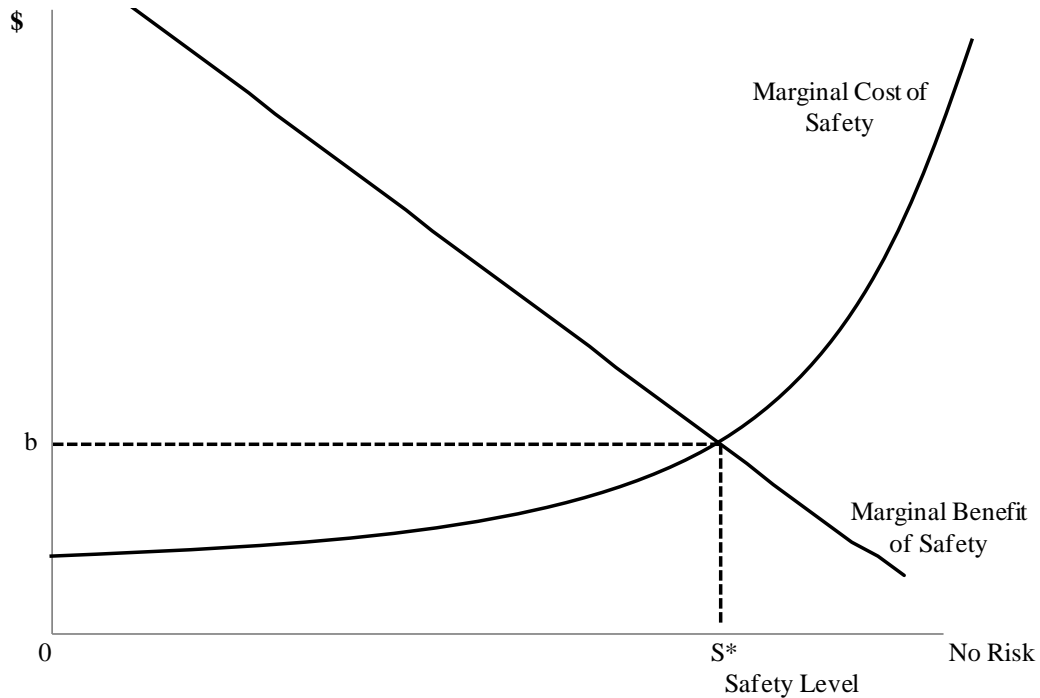
sessions, protective clothing and other safety gear must be furnished to workers, and so forth. Second, competitive pressures force many firms to operate at zero economic profits — that is, at a point at which all costs are covered and the rate of return on capital is about what it is for similar investments. Third, all of the job characteristics are already determined. The consequence of these assumptions is that if a firm undertakes a program to reduce the risk of injury, it must reduce wages to remain competitive — the more a competitive firm spends on safety, the less it can spend on other things. Although this analysis could focus on the tradeoff between safety and other working conditions, it focuses on wage tradeoffs because wages are easy to measure and form the largest component of compensation.

From the employer's perspective, there are diminishing marginal returns to safety expenditures. The first expenditure of the firm to reduce risk will have a relatively high return because the firm will clearly choose to attack the most obvious and cheaply eliminated hazard. Thus, at first, the risk and accompanying cost reductions are relatively large and the firm need not reduce wages very much to keep profits constant. At some point, additional expenditures by the firm to reduce risk have diminishing returns because all of the easy ways to solve safety problems have been used. Thus, expanding the safety effort beyond this point results in large wage reductions because further increases in safety are very costly. Furthermore, employers differ in the ease with which they can reduce hazards. In firms where injuries are costly to reduce, large wage reductions are required to keep profits constant when a safety program is implemented.

The market matches employers and employees based on, among other things, their wage-risk preference tradeoffs. Of course, not all firms are the same, and those with a comparative advantage at risk reduction have a comparative advantage at attracting relatively more risk averse workers. Similarly, not all workers are the same, and those who are relatively less risk averse will seek out firms that pay a risk premium because of their comparative disadvantage in reducing risk.

Figure VI-5 illustrates the market determination of workplace safety. The range of possible safety levels, from zero safety to no risk, is on the horizontal axis. The marginal costs to the employer of providing additional units of safety increase as the workplace becomes safer. The marginal benefits to workers of "consuming" additional units of workplace safety decline as the level of risk approaches zero. The price that workers are willing to pay for increased safety by accepting a lower wage declines as risk declines. Employers respond to workers' preferences for safety by providing additional units of safety up to the level where the marginal benefit is equal to the marginal cost. In Figure VI-5, the market determined level of workplace safety is  $S^*$ . Note that this level is to the left of the no-risk level of safety. At  $S^*$ , the employer spends the shaded area on workplace safety. The employer does not incur the additional expenditures beyond  $S^*$  because it costs more to provide the additional safety than the employees are willing to pay by accepting lower wages.

**Figure VI-5. Market Determination of Workplace Safety**



#### **4. Information, Risk, and Price Adjustments**

In order for the market to adjust to reflect risky products or risky jobs, it is necessary for the buyers in the market place to have information regarding the risk they are undertaking. If the risk is not readily observable and is not known to other people in the market place, the market price will not reflect the risk. In some instances, the manufacturers of the product may be aware of a risk of which they have not informed consumers. Certainly, many products liability cases have involved allegations of this type of situation. Asymmetric information is often reduced through regulations requiring warning labels to be placed on various products to alert consumers of the risks they are undertaking. The federal government uses a combination of disclosure regulation and explicit risk reduction regulation to deal with various kinds of risk. In general, disclosure regulation covers many financial and other economic risks, while direct regulation covers many health and safety risks.

#### **D. Tort Law**

The law of torts provides a set of legal rules that allows parties injured as the result of the actions of others to collect damages under certain well-defined circumstances. Bearing the costs of injurious acts provides an incentive to potentially liable parties to alter their behavior to reduce the risk of injury. For instance, many state courts have held grocery store owners liable to customers mugged in unlit parking lots. Impliedly, courts are saying that it is cheaper for stores to light their parking lots to prevent muggings than for customers to hire security guards or provide some other form of personal security. Several often-stated economic goals of tort law are

reviewed in this section.

### **1. Behavior Modification and Minimizing the Costs of Accidents**

The economic analysis of tort law begins with the assumption that alternative tort liability rules affect the behavior of individuals and firms in predictable ways. The problem addressed by alternative tort liability rules is how to induce both the potential tortfeasor and the potential victim to exercise an appropriate level of care. One way to characterize economic analysis of tort law is that it is primarily concerned with preventing accidents. However, within this view, it is clear that not all accidents are worth preventing. Thus, the economic perspective on tort law has concentrated on developing rules that minimize the total cost associated with accidents, where the total cost is defined as the sum of the cost of prevention and the actual cost of the accidents that occur. When the expected costs of an accident are greater than the cost of accident avoidance, then an efficient legal rule would be one that imposes liability on the party who could have avoided the accident at the lowest possible cost.

The behavior modification goal of tort law is most evident in the comparison of incentives to invest in accident avoidance activities under alternative liability rules. The starting point is an analysis of the duties owed by a reasonable person under a negligence liability standard. Four elements are traditionally required in a cause of action for negligence: (1) a duty to protect the plaintiff against unreasonable risks, (2) a failure to perform the duty, (3) actual loss or damage to the plaintiff, and (4) a reasonably close connection between defendant's conduct and the plaintiff's injury — also known as causation. All four of the elements are important because a cause of action usually fails if any one of the elements is lacking. However, the second element has received a great deal of attention due to Judge Learned Hand's famous economic formulation of the reasonable person standard.

#### **United States v. Carroll Towing Co.**

United States Court of Appeals for the Second Circuit  
159 F.2d 169 (1947)

L. HAND, Circuit Judge.

These appeals concern the sinking of the barge, "Anna C," on January 4, 1944, off Pier 51, North River. The Connors Marine Co., Inc., was the owner of the barge, which the Pennsylvania Railroad Company had chartered; the Grace Line, Inc., was the charterer of the tug, "Carroll," of which the Carroll Towing Co., Inc., was the owner. The decree in the limitation proceeding held the Carroll Company liable to the United States for the loss of the barge's cargo of flour, and to the Pennsylvania Railroad Company, for expenses in salvaging the cargo and barge; and it held the Carroll Company also liable to the Connors Company for one half the damage to the barge; these liabilities being all subject to limitation. The decree in the libel suit held the Grace Line primarily liable for the other half of the damage to the barge, and for any part of the first half, not recovered against the Carroll Company because of limitation of liability; it also held the Pennsylvania Railroad secondarily liable for the same amount that the Grace Line was liable. The Carroll Company and the Pennsylvania Railroad Company have filed assignments of error.

The facts, as the judge found them, were as follows. On June 20, 1943, the Connors Company chartered the barge, "Anna C." to the Pennsylvania Railroad Company at a stated hire per diem, by a charter of the kind usual in the Harbor, which included the services of a bargee, apparently limited to the hours 8 A.M. to 4 P.M. On January 2, 1944, the barge, which had lifted

the cargo of flour, was made fast off the end of Pier 58 on the Manhattan side of the North River, whence she was later shifted to Pier 52. At some time not disclosed, five other barges were moored outside her, extending into the river; her lines to the pier were not then strengthened. At the end of the next pier north (called the Public Pier), lay four barges; and a line had been made fast from the outermost of these to the fourth barge of the tier hanging to Pier 52. The purpose of this line is not entirely apparent, and in any event it obstructed entrance into the slip between the two piers of barges. The Grace Line, which had chartered the tug, "Carroll," sent her down to the locus in quo to "drill" out one of the barges which lay at the end of the Public Pier; and in order to do so it was necessary to throw off the line between the two tiers. On board the "Carroll" at the time were not only her master, but a "harbormaster" employed by the Grace Line. Before throwing off the line between the two tiers, the "Carroll" nosed up against the outer barge of the tier lying off Pier 52, ran a line from her own stem to the middle bit of that barge, and kept working her engines "slow ahead" against the ebb tide which was making at that time. The captain of the "Carroll" put a deckhand and the "harbormaster" on the barges, told them to throw off the line which barred the entrance to the slip; but, before doing so, to make sure that the tier on Pier 52 was safely moored, as there was a strong northerly wind blowing down the river. The "harbormaster" and the deckhand went aboard the barges and readjusted all the fasts to their satisfaction, including those from the "Anna C." to the pier.

After doing so, they threw off the line between the two tiers and again boarded the "Carroll," which backed away from the outside barge, preparatory to "drilling" out the barge she was after in the tier off the Public Pier. She had only got about seventy-five feet away when the tier off Pier 52 broke adrift because the fasts from the "Anna C," either rendered, or carried away. The tide and wind carried down the six barges, still holding together, until the "Anna C" fetched up against a tanker, lying on the north side of the pier below — Pier 51 — whose propeller broke a hole in her at or near her bottom. Shortly thereafter: i.e., at about 2:15 P.M., she careened, dumped her cargo of flour and sank. The tug, "Grace," owned by the Grace Line, and the "Carroll," came to the help of the flotilla after it broke loose; and, as both had syphon pumps on board, they could have kept the "Anna C" afloat, had they learned of her condition; but the bargee had left her on the evening before, and nobody was on board to observe that she was leaking. The Grace Line wishes to exonerate itself from all liability because the "harbormaster" was not authorized to pass on the sufficiency of the fasts of the "Anna C" which held the tier to Pier 52; the Carroll Company wishes to charge the Grace Line with the entire liability because the "harbormaster" was given an over-all authority. Both wish to charge the "Anna C" with a share of all her damages, or at least with so much as resulted from her sinking. The Pennsylvania Railroad Company also wishes to hold the barge liable. The Connors Company wishes the decrees to be affirmed.

\* \* \*

We cannot . . . excuse the Connors Company for the bargee's failure to care for the barge, and we think that this prevents full recovery. First as to the facts. As we have said, the deckhand and the "harbormaster" jointly undertook to pass upon the "Anna C's" fasts to the pier; and even though we assume that the bargee was responsible for his fasts after the other barges were added outside, there is not the slightest ground for saying that the deckhand and the "harbormaster" would have paid any attention to any protest which he might have made, had he been there. We do not therefore attribute it as in any degree a fault of the "Anna C" that the flotilla broke adrift. Hence she may recover in full against the Carroll Company and the Grace Line for any injury she suffered from the contact with the tanker's propeller, which we shall speak of as the

"collision damages." On the other hand, if the bargee had been on board, and had done his duty to his employer, he would have gone below at once, examined the injury, and called for help from the "Carroll" and the Grace Line tug. Moreover, it is clear that these tugs could have kept the barge afloat, until they had safely beached her, and saved her cargo. This would have avoided what we shall call the "sinking damages." Thus, if it was a failure in the Conner Company's proper care of its own barge, for the bargee to be absent, the company can recover only one third of the "sinking" damages from the Carroll Company and one third from the Grace Line. For this reason the question arises whether a barge owner is slack in the care of his barge if the bargee is absent.

\* \* \*

. . . [T]here is no general rule to determine when the absence of a bargee or other attendant will make the owner of the barge liable for injuries to other vessels if she breaks away from her moorings. However, in any cases where he would be so liable for injuries to others obviously he must reduce his damages proportionately, if the injury is to his own barge. It becomes apparent why there can be no such general rule, when we consider the grounds for such a liability. Since there are occasions when every vessel will break from her moorings, and since, if she does, she becomes a menace to those about her; the owner's duty, as in other similar situations, to provide against resulting injuries is a function of three variables: (1) The probability that she will break away; (2) the gravity of the resulting injury, if she does; (3) the burden of adequate precautions. Possibly it serves to bring this notion into relief to state it in algebraic terms: if the probability be called P; the injury, L; and the burden, B; liability depends upon whether B is less than L multiplied by P: i.e., whether B less than PL. Applied to the situation at bar, the likelihood that a barge will break from her fasts and the damage she will do, vary with the place and time; for example, if a storm threatens, the danger is greater; so it is, if she is in a crowded harbor where moored barges are constantly being shifted about. On the other hand, the barge must not be the bargee's prison, even though he lives aboard; he must go ashore at times. We need not say whether, even in such crowded waters as New York Harbor a bargee must be aboard at night at all; it may be that the custom is otherwise . . . ; and that, if so, the situation is one where custom should control. We leave that question open; but we hold that it is not in all cases a sufficient answer to a bargee's absence without excuse, during working hours, that he has properly made fast his barge to a pier, when he leaves her. In the case at bar the bargee left at five o'clock in the afternoon of January 3rd, and the flotilla broke away at about two o'clock in the afternoon of the following day, twenty-one hours afterwards. The bargee had been away all the time, and we hold that his fabricated story was affirmative evidence that he had no excuse for his absence. At the locus in quo — especially during the short January days and in the full tide of war activity — barges were being constantly 'drilled' in and out. Certainly it was not beyond reasonable expectation that, with the inevitable haste and bustle, the work might not be done with adequate care. In such circumstances we hold — and it is all that we do hold — that it was a fair requirement that the Conners Company should have a bargee aboard (unless he had some excuse for his absence), during the working hours of daylight.

\* \* \*

### **Notes and Questions**

**1. The Hand Formula:** Judge Hand's negligence formula is really just a cost-benefit analysis. Under the famous BPL formulation, B is the cost of undertaking a particular precautionary measure. P is the probability that an accident will occur if a particular precautionary measure is not taken. L is the loss that will result if the accident occurs. Thus, P x



L is the benefit of undertaking a particular precautionary measure because it is the weighted loss of the accident that is avoided by the precautionary measure. If  $B < PL$ , the benefit exceeds the cost and the reasonable person would have taken the precautionary measure. If they did not take this precaution, then they should be held negligent. If  $B > PL$ , the cost exceeds the benefit and the reasonable person would not have taken the precautionary measure to avoid the accident and, thus, would not be held liable for negligence if the injury occurred. What assumptions regarding risk preferences does the Hand Formula make?

**2. Hand Formula Example:** Suppose that the sunken barge in *U.S. v Carroll Towing* and its cargo are worth \$100,000. Assume that the probability that the barge would break loose if the bargee is not present is 0.0005. Paying the bargee to stay on the barge will cost the barge owner \$25. If the barge owner does not incur this \$25 expense, is his behavior negligent under the Hand rule?

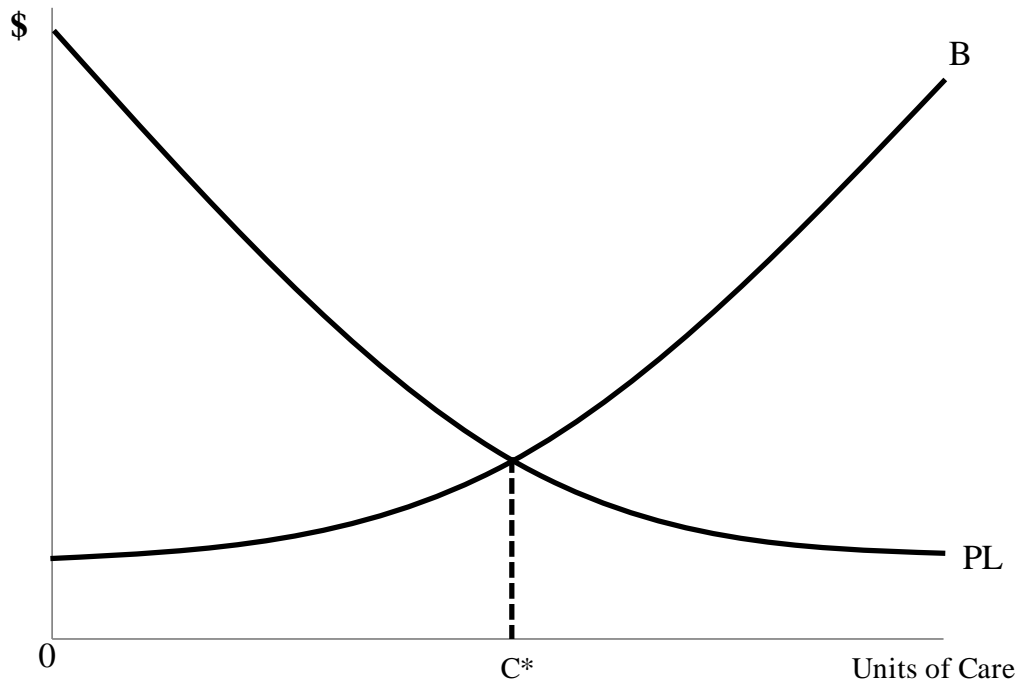
**3. Ex ante Analysis:** Because the economic model is based on behavior modification and liability rules are supposed to be a guide to behavior, the BPL test should be applied to conditions existing before the accident — at the time the potential tortfeasor was making the decision about how much to invest in accident avoidance. Thus, ex ante, the BPL formulation provides a guideline for avoiding negligence through investment in accident avoidance. Obviously, the type of activity will dictate the degree of care that is required. Under the Hand Formula, a dangerous activity which has a high probability of causing very serious injuries to many people would require great expenditures in order to avoid a judgment of negligence. On the other hand, if the activity is "safe" — that is, the probability of an accident is very low and the expected injuries are minor — then the expenditures required to avoid a judgment of negligence would be small. Under the Hand Formula, it is not negligence if an injury occurs for which the cost of preventing the injury exceeds the *expected* benefits derived from such expenditures. For example, preventing *all* injuries from flying baseballs in a stadium would be prohibitively expensive, and the likelihood and severity of injury in certain areas of the stands may be quite low. Therefore, failure to protect fans in the right field stands, where the risk of being hit by a baseball is low, may not be negligent. Failure to protect fans behind home plate, however, where the risk of injury is great, may be negligent.

**4. Least Cost Avoider of Accidents:** The behavioral modification or deterrence view of tort law generates a public policy prescription based on economic efficiency. If the goal of tort law is to minimize the total cost of accidents, then the "least cost avoider" should be held liable. Of course, this broad public policy view raises important questions about who was the least cost avoider. Alternative liability rules, such as negligence, contributory negligence, comparative negligence, and strict liability, can be viewed as placing the risk on the parties who are the least cost risk avoiders under different circumstances. In practice, the identification of the "least-cost avoider" is not always obvious. According to Professor (now-judge) Guido Calabresi, one of the leading commentators on tort law, a number of guidelines can be used to identify the party who should bear the loss of accident. According to Calabresi, the person who should bear the loss of accident should be the party who can: (1) better evaluate the risk involved; (2) better evaluate the accident-proneness of potential parties on the other side; (3) cause prices to reflect this knowledge; (4) insure most cheaply against liability; and/or (5) more likely avoid having the loss shifted in a way that reduces the incentive to avoid the loss. Guido Calabresi, *Faults, Accidents, and the Wonderful World of Blum and Kalven*, 75 *Yale Law Journal* 216 (1965). Can you determine from Calabresi's criteria what he views as the purpose (or purposes) of tort law?

**5. Marginal Analysis:** Recall from Chapter I that our basic decision rule for individuals,

firms, or government regulators is that it is optimal to engage in an activity up to the point where the marginal benefit equals the marginal cost. In this regard, a more precise statement of the Hand Formula should be based on marginal costs and marginal benefits from investing in care. This can be illustrated diagrammatically as shown in Figure VI-6. The horizontal axis indicates units of care in terms of risk avoidance activity. The B curve starts low at low units of care and rises rapidly taking on an upward sloping shape because it is assumed that the cost of reducing accidents or reducing the damages associated with accidents goes up as the likelihood of accidents or the cost of accidents is reduced — that is, on the margin, the cost of further reductions in risk increase as the units of care increase. The PL curve is downward sloping based on the notion that as more units of care are incurred, either P will be reduced because of a lower probability of accidents or L will be reduced in the sense that the damage that will be caused when more units of care are invested is likely to be less. B can be thought of as a marginal cost curve which increases with the number of units of care. The PL curve is a marginal benefit curve which suggests that the marginal benefit of additional units of care declines as more and more units of care are used. The marginal benefit and marginal cost of investing in care are equal at  $C^*$  units of care. That is the efficient standard of care. Any additional investments in care mean that the marginal cost is greater than the marginal benefit, so that the investments are not cost effective. Fewer investments in care indicate that the reduced marginal benefit from reducing care is greater than the marginal costs that are saved from reducing the care, so that making the investments would be cost effective. Under this marginal formulation of the Hand Formula, an individual should be held negligent if they failed to invest  $C^*$  units of care. That is, to the left of  $C^*$ ,  $B < PL$  and they should be held liable under the Hand negligence standard. Firms or individuals who invest more than  $C^*$  are operating in the range where  $B > PL$ , and they should not be held liable under the negligence standard. Thus, the negligence standard of care is expected to induce the efficient amount of accident avoidance expenditures.

**Figure VI-6. The Hand Formula**



**6. Strict Liability and Deterrence Effects:** An important insight derived from this model of accident avoidance is that if the liability rule is changed from negligence to strict liability — where the tortfeasor pays even if not at fault — then the injurer will theoretically invest the same amount in accident avoidance as under the negligence rule where negligence is determined by the Hand Formula. This is because up until the  $C^*$  level of care the expected cost to the injurer associated with accidents is greater than the expected cost of risk avoidance ( $B$ ). Because of those expectations, the benefit of avoiding accidents is greater than the cost of accidents up to the level  $C^*$ . Beyond that level, the cost of avoiding accidents is greater than the expected payout to injured parties. Thus, the injurer will decide not to make investments beyond  $C^*$  even though the injurer realizes that he or she will have to pay damages in the event an accident occurs. Thus, rather than focusing on deterrence and behavior modification, the strict liability standard appears to some extent to be driven by distributional considerations.

**7. Strict Liability and Victims' Incentives:** In many situations, both potential tortfeasors and potential victims can take precautionary measures to avoid accidents. Although both a strict liability rule and a negligence rule should induce tortfeasors to invest the same amount in accident avoidance, the liability rules produce differing incentives for victims. Under a rule of strict liability, potential victims have little incentive to take precautionary measures because they are compensated for any accidents (assuming the courts award damages that fully compensate victims for their losses). Under a rule of negligence, on the other hand, tortfeasors meeting the legal standard of care will not be found liable and victims will bear the loss resulting from accidents. Thus, under a negligence rule, victims have incentives to take cost-effective precautions to reduce the risk of accidents that they would receive no compensation for if

tortfeasors were found not liable. As a result, in situations where both potential tortfeasors and potential victims can take precautions to reduce the risk of accidents, negligence rules give both parties the incentive to do so.

Can you offer an economic justification for why the owner of a dog is liable for the harm it causes due to his negligence (at least initially), but the owner of a tiger is strictly liable for any harm that it causes?

**8. Defenses to Negligence:** Because the economic perspective on tort law concentrates on minimizing the total costs associated with accidents, the Hand Formula should also be applied to the behavior of potential victims. The defenses to negligence rely in part on a similar analysis of the cost of avoiding the injury and the expected loss from the injury. A potential defendant did not exercise reasonable care if the injury could have been avoided at a cost that is lower than the expected cost of the injury. A similar analysis can be applied to the plaintiff's decision making. A plaintiff did not exercise reasonable care if she failed to take precautions that cost less than the expected cost of the injury.

**9. Contributory and Comparative Negligence:** Contributory and comparative negligence are defenses to negligence when the plaintiff is at least partially responsible for his or her injury. The defense of **contributory negligence** is the failure of the plaintiff — the injured party — to exercise reasonable care that contributes to the injury. Traditionally, contributory negligence would completely bar recovery for injuries. Most states have now adopted the doctrine of **comparative negligence**, which allows a proration of the damages resulting from the combined negligence of the parties. If a plaintiff's injuries amount to \$100,000, and they are 40% the result of plaintiff's negligence and 60% the result of defendant's negligence, then the plaintiff's recovery would be limited to only \$60,000. Compared to a simple negligence rule with no defenses, do either contributory negligence or comparative negligence defenses change potential victims' incentives to take precautionary measures to avoid accidents? Do they change potential tortfeasors' incentives? The answer is much more complicated than it seems. See, for example, Oren Bar-Gill & Omri Ben-Shahar, *The Uneasy Case for Comparative Negligence*, 5 *American Law and Economics Review* 433 (2003).

What about an accident between a Kia and a Ferrari? Assuming both drivers are equally negligent, is it fair to expect the Kia driver to be responsible for 50% of the damages to the Ferrari? In a system of strict contributory negligence, both drivers would be responsible for the damages to their own cars, because their own negligence would prevent recovery in tort. Does comparative negligence lead to the externalization of costs associated with owning expensive cars? Do the rich benefit more from comparative negligence? See Michael Krauss, *In Defense of Contributory Negligence*, PointofLaw.com (Oct. 31, 2008), <http://www.pointoflaw.com/archives/2008/10/in-defense-of-c.php>.

**10. Negligence Versus Strict Liability, and Activity Levels:** Professor Steven Shavell has added some additional elements to the basic BPL model of negligence liability:

By definition, under the negligence rule all that an injurer needs to do to avoid the possibility of liability is to make sure to exercise due care if he engages in his activity. Consequently, he will not be motivated to consider the effect on accident losses of his choice of whether to engage in his activity or, more generally, of the level at which to engage in his activity; he will choose his level of activity in accordance only with the personal benefits so derived. But surely any increase in his level of activity will typically raise expected accident losses (holding constant the level of care). Thus he will be led to choose too high a level of activity; the negligence rule is not "efficient."

Consider by way of illustration the problem of pedestrian-automobile accidents (and, . . . let us imagine the behavior of pedestrians to be fixed). Suppose that drivers of automobiles find it in their interest to adhere to the standard of due care but that the possibility of accidents is not thereby eliminated. Then, in deciding how much to drive, they will contemplate only the enjoyment they get from doing so. Because (as they exercise due care) they will not be liable for harms suffered by pedestrians, drivers will not take into account that going more miles will mean a higher expected number of accidents. Hence, they will do too much driving; an individual will, for example, decide to go for a drive on a mere whim despite the imposition of a positive expected cost to pedestrians.

However, under a rule of strict liability, the situation is different. Because an injurer must pay for losses whenever he is involved in an accident, he will be induced to consider the effect on accident losses of both his level of care and his activity level. His decisions will therefore be efficient. Because drivers will be liable for losses sustained by pedestrians, they will decide not only to exercise due care in driving, but also to drive only when the utility gained from it outweighs expected liability payments to pedestrians. Steven Shavell, *Strict Liability versus Negligence*, 9 *Journal of Legal Studies* 1, 2–3 (1980). In effect, a strict liability standard forces the manufacturer to be an insurer of unavoidable risks. Thus, according to Shavell, a strict liability standard forces the manufacturer to contemplate the level of unavoidable risk that results from its activity. Why, even in this case, is the manufacturer assumed to be the least cost avoider? Does Shavell's "externality argument" simply obscure the issue of economic efficiency?

**11. Court Errors:** Suppose a particular court systematically imposes excessive damages (greater than the actual harm resulting from accidents). What will this do to potential tortfeasors' incentives to take precautionary measures to avoid accidents under a rule of strict liability or a rule of negligence? What if the errors are random (they are not systematically too high or too low)?

Suppose instead that the court systematically sets the legal standard of care above the efficient level of care ( $C^*$  in Figure VI-6). What will this do to potential tortfeasors' incentives to take precautionary measures to avoid accidents under a rule of negligence? What if the errors are random (they are not systematically above or below the efficient level of care)?

**12. Behavior Modification and Compensation:** The pure behavioral modification perspective on tort law suggests that compensation of the accident victim is irrelevant. The important point is that the potential tortfeasor expects to pay damages if held liable and, thus, responds to the incentives created by the liability rule. Damages can be paid to anyone — the victim, the state, the authors of this casebook — because it is assumed that the potential recipient of the damages does not affect the potential tortfeasor's incentives. After all, distributional considerations are routinely ignored when the tortfeasor is not liable in the sense that the victim isn't compensated for the injury suffered. This is not to suggest, however, that deterrence should be the exclusive goal of tort law.

## 2. Loss Spreading and Insurance

Another economic approach to tort law is to consider it as a form of insurance. Recall that risk averse individuals buy insurance to reduce the downside risk associated with many of life's activities. In the absence of tort liability, tort victims would be forced to bear large losses. Tort rules can be used to compensate unlucky victims. According to this view, the party who is

in the best position to spread the loss of the injury should be held liable. With respect to injuries resulting from products, for example, it is argued that the manufacturer is often in the best position to cover the costs of compensation because a portion of the expected accident costs is reflected in the price of the product. All consumers, not just the unlucky few who happen to be injured by the product, bear some of the costs. All purchasers of a product pay a slightly higher price that, in effect, buys them an insurance policy for compensation in the event of injury while using the product.

**Greenman v. Yuba Power Products, Inc.**

Supreme Court of California

377 P.2d 897 (1963)

TRAYNOR, J.

Plaintiff brought this action for damages against the retailer and the manufacturer of a Shopsmith, a combination power tool that could be used as a saw, drill, and wood lathe. He saw a Shopsmith demonstrated by the retailer and studied a brochure prepared by the manufacturer. He decided he wanted a Shopsmith for his home workshop, and his wife bought and gave him one for Christmas in 1955. In 1957 he bought the necessary attachments to use the Shopsmith as a lathe for turning a large piece of wood he wished to make into a chalice. After he had worked on the piece of wood several times without difficulty, it suddenly flew out of the machine and struck him on the forehead, inflicting serious injuries. About 10 1/2 months later, he gave the retailer and the manufacturer written notice of claimed breaches of warranties and filed a complaint against them alleging such breaches and negligence.

After a trial before a jury, the court ruled that there was no evidence that the retailer was negligent or had breached any express warranty and that the manufacturer was not liable for the breach of any implied warranty. Accordingly, it submitted to the jury only the cause of action alleging breach of implied warranties against the retailer and the causes of action alleging negligence and breach of express warranties against the manufacturer. The jury returned a verdict for the retailer against plaintiff and for plaintiff against the manufacturer in the amount of \$65,000. The trial court denied the manufacturer's motion for a new trial and entered judgment on the verdict. The manufacturer and plaintiff appeal. Plaintiff seeks a reversal of the part of the judgment in favor of the retailer, however, only in the event that the part of the judgment against the manufacturer is reversed.

Plaintiff introduced substantial evidence that his injuries were caused by defective design and construction of the Shopsmith. His expert witnesses testified that inadequate set screws were used to hold parts of the machine together so that normal vibration caused the tailstock of the lathe to move away from the piece of wood being turned permitting it to fly out of the lathe. They also testified that there were other more positive ways of fastening the parts of the machine together, the use of which would have prevented the accident. The jury could therefore reasonably have concluded that the manufacturer negligently constructed the Shopsmith. The jury could also reasonably have concluded that statements in the manufacturer's brochure were untrue, that they constituted express warranties, and that plaintiff's injuries were caused by their breach.

The manufacturer contends, however, that plaintiff did not give it notice of breach of warranty within a reasonable time and that therefore his cause of action for breach of warranty is barred by section 1769 of the Civil Code. . . .

\* \* \*

The notice requirement of section 1769, however, is not an appropriate one for the court to adopt in actions by injured consumers against manufacturers with whom they have not dealt. "As between the immediate parties to the sale [the notice requirement] is a sound commercial rule, designed to protect the seller against unduly delayed claims for damages. As applied to personal injuries, and notice to a remote seller, it becomes a booby-trap for the unwary. The injured consumer is seldom 'steeped in the business practice which justifies the rule,' and at least until he has had legal advice it will not occur to him to give notice to one with whom he has had no dealings." . . . We conclude, therefore, that even if plaintiff did not give timely notice of breach of warranty to the manufacturer, his cause of action based on the representations contained in the brochure was not barred.

Moreover, to impose strict liability on the manufacturer under the circumstances of this case, it was not necessary for plaintiff to establish an express warranty. A manufacturer is strictly liable in tort when an article he places on the market, knowing that it is to be used without inspection for defects, proves to have a defect that causes injury to a human being. Recognized first in the case of unwholesome food products, such liability has now been extended to a variety of other products that create as great or greater hazards if defective.

Although in these cases strict liability has usually been based on the theory of an express or implied warranty running from the manufacturer to the plaintiff, the abandonment of the requirement of a contract between them, the recognition that the liability is not assumed by agreement but imposed by law, and the refusal to permit the manufacturer to define the scope of its own responsibility for defective products make clear that the liability is not one governed by the law of contract warranties but by the law of strict liability in tort. Accordingly, rules defining and governing warranties that were developed to meet the needs of commercial transactions cannot properly be invoked to govern the manufacturer's liability to those injured by its defective products unless those rules also serve the purposes for which such liability is imposed.

We need not re canvass the reasons for imposing strict liability on the manufacturer. They have been fully articulated in [our prior cases]. The purpose of such liability is to insure that the costs of injuries resulting from defective products are borne by the manufacturers that put such products on the market rather than by the injured persons who are powerless to protect themselves. Sales warranties serve this purpose fitfully at best. In the present case, for example, plaintiff was able to plead and prove an express warranty only because he read and relied on the representations of the Shopsmith's ruggedness contained in the manufacturer's brochure. Implicit in the machine's presence on the market, however, was a representation that it would safely do the jobs for which it was built. Under these circumstances, it should not be controlling whether plaintiff selected the machine because of the statements in the brochure, or because of the machine's own appearance of excellence that belied the defect lurking beneath the surface, or because he merely assumed that it would safely do the jobs it was built to do. It should not be controlling whether the details of the sales from manufacturer to retailer and from retailer to plaintiff's wife were such that one or more of the implied warranties of the sales act arose. "The remedies of injured consumers ought not to be made to depend upon the intricacies of the law of sales." To establish the manufacturer's liability it was sufficient that plaintiff proved that he was injured while using the Shopsmith in a way it was intended to be used as a result of a defect in design and manufacture of which plaintiff was not aware that made the Shopsmith unsafe for its intended use.

\* \* \*

The judgment is affirmed.

## Notes and Questions

**1. *Overlap of Deterrence and Loss Spreading Goals:*** It is often argued that the manufacturer's superior ability to spread the loss can be consistent with the goal of minimizing the costs of accidents, as Justice Traynor wrote in *Escola v. Coca Cola Bottling Co.*, 150 P.2d 436 (1944):

. . . In my opinion it should now be recognized that a manufacturer incurs an absolute liability when an article that he has placed on the market, knowing that it is to be used without inspection, proves to have a defect that causes injury to human beings. . . . Even if there is no negligence, . . . public policy demands that responsibility be fixed wherever it will most effectively reduce the hazards to life and health inherent in defective products that reach the market. It is evident that the manufacturer can anticipate some hazards and guard against the recurrence of others, as the public cannot. Those who suffer injury from defective products are unprepared to meet its consequences. The cost of an injury and the loss of time or health may be an overwhelming misfortune to the person injured, and a needless one, for the risk of injury can be insured by the manufacturer and distributed among the public as a cost of doing business. It is to the public interest to discourage the marketing of products having defects that are a menace to the public. If such products nevertheless find their way into the market it is to the public interest to place the responsibility for whatever injury they may cause upon the manufacturer, who, even if he is not negligent in the manufacture of the product, is responsible for its reaching the market. However intermittently such injuries may occur and however haphazardly they may strike, the risk of their occurrence is a constant risk and a general one. Against such a risk there should be general and constant protection and the manufacturer is best situated to afford such protection.

*Id.* at 453 (Traynor, J., concurring). This view, in effect, turns a manufacturer into an insurer. Justice Traynor appears to be assuming that manufacturers can insure against non-negligent risk more cheaply than consumers. Is there any justification for this assumption?

**2. *The Limits of Risk Aversion and Loss Spreading:*** Although loss spreading is a widely accepted goal of tort law, some judges do recognize that there must be limits to how many injuries can be compensated under such a theory. In *Shepard v. Superior Court*, 76 Cal. App. 3d 16 (1977), a dissenting judge argued against allowing recovery for the plaintiffs' alleged physical injuries resulting from the emotional shock of watching their daughter being killed in an automobile accident:

. . . To start with, it is noted that the avowed purpose of imposing strict liability upon the manufacturer is twofold: (1) loss-distribution or risk-spreading and (2) injury-reduction by enhanced safety. The first rationale, risk-spreading, holds the manufacturer liable for injuries resulting from the use of his product because he is in the best position to distribute the loss either by insurance or by increasing the price of his product. . . . The second rationale, the theory of injury reduction, holds the manufacturer liable because he is in the best position to discover and correct the dangerous aspects of his products before any injury occurs. Again, the manufacturer may pass on to the consumer the increased product costs by incorporating them in the purchase price of the merchandise.

Although since its inception the courts have generally tended to broaden the scope of products liability, there are few cases, if any, which have embarked on a thorough and delicate analysis to explore whether the above stated policy goals are indeed promoted by



the ever-expanding scope of enterprise liability. It is time for such an examination.

The basic facts of economy teach us that the fashionable trend of a wholesale extension of strict liability proves to be counterproductive in many instances by hampering and arresting, rather than promoting, the policy objectives underpinning the doctrine. . . . While some portion of the ever growing safety and insurance cost may pass directly to the consumer by way of a higher dollar price, the remainder will take the form of decreased quality not affecting safety and decreased profits. The decreased profits affect the manufacturers first (among them mainly the large segment of small businessmen with limited or marginal capital who have to shoulder the strict enterprise liability side by side with the huge corporations), then society as a whole. The motion and realistic operation of economic forces have been graphically described by one observer as follows: "*Decreased profits, however, do not stop with the manufacturer. He distributes them to the shareholders of his corporation, just as he distributes increased prices to the consumers of his product. Moreover, decreased profits do not stop with the shareholders. Rather, in more or less attenuated form, they pass on to other, broader classes. The major distribution of decreased profits occurs when shareholders switch their investment to other, more profitable enterprises. When this happens, the liability-bearing manufacturer's enterprise loses its ability to attract investment capital resulting in decreased industrial activity. This decreased activity results in losses to several categories. First, the consumer will feel the loss because the manufacturer's ability to produce a better, safer product will diminish. Second, reduced industrial activity will affect labor. Severely diminished profits may force the manufacturer out of business. Even less drastic reductions, however, could reduce the number of new jobs. Finally, reduced economic activity will affect the entire society, in a more or less attenuated form, through lower tax revenues, lower wages, and lower profits for distribution.*" (Alden D. Holford, *The Limits of Strict Liability for Product Design and Manufacture* (1973) 52 *Tex.L.Rev.* 81, 87, emphasis added.)

Paying heed to economic realities rather than our own fancy, the courts as a matter of judicial policy must stop the further extension of the strict liability of entrepreneurs, at least to areas where, as here, the determination of damages is speculative and conjectural rather than real and definable. In doing so, we are in line with established law which holds that the manufacturer is not an insurer of the product and that the strict liability of entrepreneurs may not be equated with absolute, limitless liability. As has been emphasized time and time again, in determining the parameters of enterprise liability we must draw a proper balance between the need for adequate recovery and the survival of viable enterprises. The guiding principles to achieve these goals are judicial temperance, evenhandedness and, first and foremost, fairness to all.

*Id.* at 26–31. This opinion clearly recognizes the presence of opportunity costs in a loss spreading theory of strict liability. Should courts consider such costs in their application of liability rules? Does Justice Traynor consider such costs in *Greenman*?

### **The Current Insurance Crisis and Modern Tort Law**

George Priest

96 *Yale Law Journal* 1521 (1987)

This paper is an effort to understand the source of the crisis in insurance that has recently disrupted product and service markets in the United States. From press accounts, the crisis

seemed to peak in the early months of 1986, when reports became common of extraordinary changes in commercial casualty insurance markets. Insurers had increased premiums drastically for an unusual set of products, such as vaccines, general aircraft, and sports equipment, and for an equally diverse set of services, such as obstetrics, ski lifts, and commercial trucking. In still other cases — intrauterine devices, wine tasting, and day care — insurers had refused to offer coverage at any premium, forcing these products and services to be withdrawn from the market.

\* \* \*

This paper argues that the characteristic of contemporary tort law most crucial to understanding the current crisis is the judicial compulsion of greater and greater levels of provider third-party insurance for victims. The progressive shift to third-party corporate insurance coverage, since its beginnings in the mid-1960's, has systematically undermined insurance markets.

\* \* \*

This explanation of the crisis uncovers what I believe to be a tragic paradox of our modern civil liability regime. The expansion of liability since the mid-1960's has been chiefly motivated by the concern of our courts to provide insurance to victims who have suffered personal injury. The most fundamental of the conceptual foundations of our modern law is that the expansion of tort liability will lead to the provision of insurance along with the sale of the product or service itself, with a portion of the insurance premium passed along in the product or service price. Expanded tort liability, thus, is a method of providing insurance to individuals, especially the poor, who have not purchased or cannot purchase insurance themselves. This insurance rationale suffuses our modern civil law, and must be acknowledged as one of the great humanitarian expressions of our time.

The paradox exposed by my theory is that the expansion of tort liability has had exactly the opposite effect. The insurance crisis demonstrates graphically that continued expansion of tort liability on insurance grounds leads to a reduction in total insurance coverage available to the society, rather than to an increase. The theory also shows that the parties most drastically affected by expanded liability and by the current insurance crisis are the low-income and poor, exactly the parties that courts had hoped most to aid.

\* \* \*

## II. MODERN TORT LAW AND ITS ECONOMIC EFFECTS

Since the early 1960's, courts have steadily expanded tort liability for injuries suffered in the context of product and service use. These changes in the law result from the acceptance of a coherent and powerful theory that justifies the use of tort law to compensate injured parties, a theory its founders called "enterprise liability." According to the theory, expanded provider liability serves three important functions: to establish incentives for injury prevention; to provide insurance for injuries that cannot be prevented; and to modulate levels of activity by internalizing costs, including injury costs.

The second feature of the theory — the importance of providing insurance for unpreventable losses — is most crucial for understanding our current insurance crisis. According to enterprise liability theory, expanded legal liability does more than achieve optimal control of accident and activity rates. Expanded tort liability improves social welfare, in addition, because it provides a form of compensation insurance to consumers. A provider, especially a corporate provider, is in a substantially better position than a consumer to obtain insurance for product- or service-related losses, because a provider can either self-insure or can enter one insurance contract covering all consumers — in comparison to the thousands of insurance contracts the set

of consumers would need — and can easily pass the proportionate insurance premium along in the product or service price. Most importantly, to tie insurance to the sale of the product or service will provide insurance coverage to consumers who might not otherwise obtain first-party coverage, in particular the poor or low-income among the consuming population.

The insurance rationale was central to the first judicial adoption of enterprise liability theory, by the California Supreme Court in *Greenman v. Yuba Power Products*, in 1963. The approach was rapidly extended across the various state jurisdictions, first in the products liability field and, later, in other areas of tort law. Briefly, however, enterprise liability theory has justified both restrictions in available legal defenses and expansion of substantive liability standards.

\* \* \*

The economic effects of steadily increasing provider liability . . . are quite simple in structure. A liability rule can compel providers of products and services to make investments that reduce the accident rate up to the level of optimal (cost-effective) investments. After providers have invested optimally in prevention, however, any further assignment of liability affects only the provision of insurance. More extensive provider liability will generate more extensive provider insurance and nothing more.

The expansion of liability under modern tort law has obviously increased the provision of provider insurance. Any standard beyond a bare cost-benefit test (often identified with negligence) will provide an insurance effect. Courts, of course, have extended liability far beyond the simple cost-benefit standard. Thus, modern tort law compels a very substantial level of provider insurance.

More precisely, modern tort law has broadly shifted the insurance obligation from first-party insurance to third-party or self-insurance by providers. Even a bare-bones cost-benefit standard has insurance consequences: Such a standard creates an obligation of potential victims to obtain market insurance or to self-insure for unpreventable losses. Modern tort law has shifted that obligation to providers, requiring providers either to obtain third-party market insurance or to self-insure for the losses suffered by consumers of their products or services. The expansion of tort liability since the mid-1960's has expanded the range of contexts in which provider insurance must be offered. Courts understand this point perfectly. Much of the modern extension of tort liability has been expressly justified by the salutary insurance consequences that are supposed to result. Thus it is a paradox that the modern regime somehow has led to the reduction of insurance availability.

\* \* \*

### III. HOW INSURANCE OPERATES

\* \* \*

Tort law, of course, provides insurance through a third-party mechanism: the insurer pays money to the victim through the medium of the product or service provider who purchases the insurance policy. Although third-party institutional arrangements are somewhat more complex than first-party arrangements, the determinants of the insurance function are the same.

In third-party insurance, there are two sets of risk pools. Consumers of products or services comprise one set of risk pools. To consumers, the insurance policy and the premium are tied to the sale of the product. It is advantageous to define risk pools narrowly for consumers in third-party contexts for exactly the same reasons that it is advantageous to define risk pools narrowly in typical first-party insurance contexts. Defining consumer risk pools narrowly increases product sales, because the premium added to the price of the product more closely

approximates the consumer's expected loss.

Manufacturers, for example, will attempt to segregate consumers into risk pools by product design and by advertising and marketing techniques. A chain-saw manufacturer, for example, may design one model appropriate for industrial use and a second model appropriate for occasional gathering of firewood. Of course, such design differences may also be related to different consumer preferences for product features — in this example, features related to safety. But the point is that, if the demand for chain saw injury insurance coverage differs between the professional and the weekender, both consumers and the manufacturer will gain if the manufacturer can design products that differentiate the two markets. In this example, market differentiation would reduce one very broad risk pool into two more narrow risk pools. Narrowing the pools allows the manufacturer to charge different insurance premiums to the two markets and to increase product sales.

Similarly, the accident insurance premium added to the price of an airline ticket from the United States to Europe will be greater on, say, the Concorde than on low-budget or charter lines. The risk of an accident among the various airlines may be the same, but the accident payout risks brought into the pool by passengers on the Concorde are likely to be much greater than the risks brought by charter passengers, if only because of their greater expected future income. The third-party insurance premium must be adjusted in response. In this respect, the qualitative differences between the Concorde and the charters in terms of accident insurance are no different than qualitative differences in meals, time of transit, or other amenities. Indeed, much of the attraction of the charters derives exactly from the ability of these firms to narrow the pool of consumers of their product. Those passengers who travel on low-fare flights are those who prefer or are willing to tolerate lower levels of amenities in return for a lower ticket price.

The second set of risk pools within third-party insurance contexts includes the service and product providers themselves. Providers of products and services purchase market insurance for the same purposes as any first-party insured: to equalize monetary returns over time in the face of some probabilistic chance of loss. Providers will choose market insurance if its costs are lower than the costs of alternatives.

The costs a provider faces by deciding not to purchase insurance depend upon diversification within the provider firm or the provider's ability to diversify risk by other means. As suggested above, marketing different models of a product is a form of diversification, if the risks of loss attending sales of the respective models are uncorrelated. Of course, organization in the conglomerate form or investing retained earnings in diversified assets are other ways in which providers can self-insure.

Insurance companies, however, can (and do) compete for the custom of providers by trying to define narrow risk pools that make market insurance more attractive than self-insurance alternatives. Insurers attempt to aggregate within a pool a set of providers whose risks are uncorrelated, and will set individual premiums for the firms according to the risk each brings to the pool. The insurer's diversification of risk, again, can be achieved either by aggregating a very disparate pool, by holding other assets whose riskiness is uncorrelated, or by reinsuring — hiring another insurer to provide meta-diversification.

These simple insurance principles seem very general, but they provide an explanation for the insurance crisis we are currently observing. The next Section attempts to apply these principles to the changes in tort law discussed in Section II, in order to predict the effects of modern law on insurance availability.

#### IV. HOW CONTEMPORARY TORT LAW AFFECTS INSURANCE MARKETS

. . . [T]he expansion of corporate liability has progressively undermined the insurance function by increasing the variance (coefficient of variation) of existing insurance risk pools. . . . [I]ncreasing the variance of a risk pool endangers the pool because it increases the likelihood that the pool will unravel as low-risk members drop out, either by self-insuring or by ceasing to engage in the potentially injury-related activity. This Section . . . shows that contemporary tort law has restricted rather than expanded insurance availability. The parties that have been most adversely affected are the low-income and poor who, in terms of tort recoveries, are the low-risk members of the consuming population.

\* \* \*

I believe that there are very clear insurance reasons why the shift towards third-party coverage has undermined the commercial casualty insurance industry, generating the crisis. In comparison to first-party insurance, third-party tort law insurance provides coverage in excessive amounts, in a manner that substantially restricts risk segregation, and at costs that far exceed the costs of first-party insurance. For both consumer and provider risk pools, these differences will increase the correlation of risks within existing pools and, as a consequence, increase the extent of adverse selection, leading to the breakdown of the pools.

\* \* \*

Provider tort law insurance coverage differs substantially from first-party insurance coverage. One of the objectives of the tort system is to create incentives for appropriate investments in preventing injury. To obtain optimal incentives for injury prevention, a party that has violated a legal standard must pay full losses to the victim, including both pecuniary and non-pecuniary losses.

The award of both pecuniary and non-pecuniary losses, however, is inappropriate for providing optimal insurance for unpreventable losses. The effort to extend insurance coverage through modern tort law represents a confusion of incentive objectives with insurance objectives. Third-party insurance payments administered through the tort system differ from first-party insurance payments in two ways. First, . . . no first-party insurance market provides coverage of non-pecuniary losses. Non-pecuniary losses do not affect the marginal value of wealth across states of the world. In addition, moral hazard and adverse selection problems make coverage of these losses exceedingly costly. Losses representing pain and suffering or other emotional effects of an injury, therefore, are never insured in first-party markets because it is not worthwhile for consumers to pay the premiums necessary to support coverage of them.

Secondly, deductibles and co-insurance are features of every first-party insurance contract. Third-party insurance through the tort system, in contrast, never incorporates deductibles or co-insurance to control victim moral hazard. Yet victim moral hazard is as serious a problem in a third-party context as in a first-party context. Preferences for extra visits to a doctor, prolonged hospitalization, or more advanced forms of medical treatment do not diminish because the source of the injury is a third-party defendant.

These two differences mean that, for the same injury, first-party insurance coverage — which corresponds to what consumers are willing to purchase — is substantially different in magnitude than the third-party insurance coverage provided through tort law.

\* \* \*

The shift from first-party to third-party insurance sources, thus, will prompt greater expenditures for advanced medical care, as well as more extended and elaborate hospitalization and subsequent care. For example, holding severity of injury constant, the frequency of claims for twenty-four hour nursing care is likely to be substantially higher under third-party tort law

insurance than under first-party insurance.

\* \* \*

Some might regard the additional level of insurance coverage provided by tort law to be beneficial to consumers because it affords them greater compensation for the injuries they suffer. But this view misunderstands the consumer interest in insurance. Of course, after an injury has been suffered, the victim would prefer a greater to a lesser award. But the ex ante interest of the victim is an award tied to the victim's pecuniary losses, not an award greater than these losses. Where the victim is the product or service purchaser, the victim must pay for the insurance in advance. To compel insurance greater than the amount demanded by the purchaser reduces, rather than increases, his or her welfare. To illustrate: if my \$100,000 home burns down, I of course would be happier if my insurer gave me \$234,000 rather than the \$100,000 of coverage I purchased. But, I would object strongly if I were compelled in advance of the fire to purchase \$234,000 coverage since I could replace the home in its entirety for an amount in the range of \$100,000. A similar concern for optimal insurance extends to contexts involving pure third-party injuries — for example, when bystanders are harmed by the product or service use of others. In modern society, all of us are at once product and service purchasers and bystanders of products and services used by others. Again, in terms solely of insurance, each of us ex ante prefers that the optimal level of insurance be provided by tort law, the level that optimizes insurance coverage subject to insurance costs.

The provision of insurance coverage through tort law in amounts greater than consumers would willingly purchase has additional effects that implicate the recent crisis. An increase of 64% to 134% in the level of insurance coverage under tort law will not operate as a scalar, but will increase consumer risk pool variance and will lead to the unraveling of consumer risk pools.

The increase in the level of insurance coverage from the shift to the third-party tort mechanism is not likely to be uniform over all cases. The empirical observation that pain and suffering awards constitute 47% of total damages is an average figure. Pain and suffering and other non-pecuniary amounts comprise a much higher proportion of large damage judgments. For this reason, risk pool variance is likely to be greater under third-party tort insurance than under first-party insurance.

More importantly, segregating risks into narrow risk pools is substantially more difficult in third-party than in first-party insurance contexts. First-party insurers, by using insurance applications, can distinguish insureds by age, income, occupation, the level of coverage desired, and other personal characteristics related to levels of risk brought to the insurance pool. Moreover, the administration of first-party insurance allows the insurer to distinguish insureds by past loss experience. The collection of these data allows a first-party insurer to define risk pools of very narrow scope, increasing the likelihood that low-risk individuals will find insurance attractive. Narrow risk differentiation maximizes the availability of insurance.

Very little information about individual risks, however, is available to third-party insurers. A product manufacturer, for example, may design and market a product with reference to characteristics of discrete sets of average consumers. But the manufacturer must sell the product on equivalent terms to all who wish to buy it, and cannot distinguish among consumers with respect to the insurance policy provided in the product price.

Some products, of course, will attract relatively homogeneous sets of consumers. A very wide range of products, however, are accessible to and are purchased by consumers of different income levels. Studies of consumers of individual products show that, for virtually all products, the income levels and personal characteristics of consumers of the product differ widely.

The difficulty of segregating risk pools in third-party contexts means that third-party tort insurance pools are likely to be substantially broader than first-party pools even without the effects, described above, of levels of coverage. Compare, for example, the risks of non-preventable injuries from auto use. The first-party insurer can create separate driver pools for teenagers and other age groups; it can segregate insureds by levels of driving, total mileage, distance from home to office, and car type; and it can rate the policies by accident experience and by moving violations within previous time periods. It can allow the insured to choose whether to purchase medical expense and disability coverage in the auto policy or to rely on separate medical and disability policy coverage set according to the deductible the insured prefers and according to the insured's income level. Each of these techniques helps keep premiums low for the low-risk drivers of the consuming population — those who drive little, are very skilled or careful, or generate small claims because of low expected income losses.

In contrast, the auto manufacturer — that must buy third-party liability insurance for all those injured in its cars and pass on the premiums in vehicle prices — can implement none of these distinctions. Some auto models may be more or less attractive to commuters, to teenagers, or to the very wealthy, but, except for these crude distinctions, the auto manufacturer must provide insurance to all who buy the model, high-risk and low-risk alike. Consequently, the variance in the insurance pool is vastly greater in the third-party context, and the premium is commensurately higher — even if the same level of coverage is offered. Of course, given the greater amount of coverage provided under a third-party policy, the variance and the premium are higher yet.

One of the most seriously deleterious effects of lumping consumers into undifferentiated third-party risk pools is glaringly inconsistent with the judicial objective of aiding the poor. That is the regressive redistributive effect of third-party insurance. The largest items of damages in most third-party personal injury contexts, especially those involving permanent disability, are lost income and pain and suffering, which are highly correlated with individuals' expected future income streams. As a consequence, these damage elements constitute the largest component of the third-party insurance premium tied to the sale of any product or service.

The third-party premium is set with reference to average expected loss. Thus, the high correlation of total awards with income means that premiums reflect the average income of the population of consumers. The implication of charging each consumer a premium related to average income is that consumers with high incomes are charged a premium lower than their expected loss, and consumers with low incomes are charged a premium higher than their expected loss. Third-party insurance thus requires low-income consumers to subsidize high-income consumers.

\* \* \*

As a consequence, the disadvantages of third-party insurance coverage are substantial. Courts justified third-party insurance coverage based on how easy it seemed to be for manufacturers or service providers to aggregate risks by adding an insurance premium to the price of the product or service. Whatever comparative advantage providers enjoy in risk aggregation, however, is overwhelmed by the disadvantages of excessive coverage, the inability to segregate risks in third-party contexts, and regressive distributive effects.

How do the differences between first-party and third-party tort insurance mechanisms affect the behavior of consumers? The shift towards greater corporate-provided tort law insurance will lead low-risk consumers to reduce consumption of products whose prices incorporate high tort insurance premiums. For low-risk consumers, especially low-income

consumers, the tort law insurance premium tied to the product or service price may be much greater than the benefit the insurance provides. As a consequence, though the effect may be subtle, these consumers will drop out of the market.

\* \* \*

#### Notes and Questions

**1. *The Enterprise Liability Paradox:*** Enterprise liability is motivated by a desire to provide insurance to victims of non-negligent accidents. At its base, the theory supposes that expanded tort liability will lead to the provision of insurance along with the sale of the product or service itself. Manufacturing and service providers will pass on as much of the liability insurance premium through price as is possible, given demand elasticity. According to the theory, the poor are especially benefitted by expanded liability because they could not otherwise afford to purchase insurance for such losses. But Professor Priest suggests that this ever-expanding liability has had a paradoxical effect. That is, enterprise liability theory has led to a reduction in the total insurance coverage available to society, rather than an increase. Particularly harmed by enterprise liability under Priest's analysis are the poor. Do you find Professor Priest's analysis persuasive? If so, what type of action should be taken to correct this judicially imposed problem?

**2. *Variance in Risk Pools:*** At the heart of Professor Priest's analysis is the fact that enterprise liability has led to an increase in risk pool variance. In general, an individual is willing to purchase insurance if the premium paid reflects the expected damages. When insurers are able to segregate risk pools, they are able to take full advantage of the law of large numbers — they can predict with certainty the total losses which any risk pool will incur. In other words, segregating risk pools and taking advantage of the law of large numbers allows insurers to substantially reduce the variance of expected losses. In turn, this reduced variance allows insurers to set premiums at a level that reflects the expected damages of risk pool members. Anything that would cause variance to increase results in the premium charged being less reflective the expected damages of some risk pool members. For high-risk members of the pool this is a subsidy. Low-risk members find that the premium is too high in comparison to expected damages. Therefore, low-risk members drop out of the risk pool. Within this line of reasoning, Professor Priest identifies two factors that have contributed to an insurance crisis. First are damages for non-pecuniary losses and the corresponding adverse selection and moral hazard problems associated with such damages. Second are the greater difficulties in risk pool segregation due to the third-party context. Can you think of other factors that contribute to risk pool variance? Does enterprise liability have any impacts that reduce risk pool variance? Are there any mitigating factors that justify a continuation of enterprise liability theory?

**3. *Ending the Vicious Circle:*** Following the reasoning of Professor Priest's analysis leads one into a vicious circle. As low-risk insureds drop out of the risk pool, premiums increase. The increased premiums cause further dropouts and so on. How does this vicious circle resolve itself? Professor Priest provides examples of several services that were forced out of the market due to enterprise liability theory. Is this result acceptable? Justifiable? What do such results portend for the future of American business?

**4. *Ultrahazardous Activities, Personal Injury, and Risk Spreading:*** Is the argument for imposition of enterprise liability more powerful when applied to ultrahazardous activities? In *Richman v. Charter Arms Corp.*, 571 F. Supp. 192 (E.D. La. 1983), a murder victim's mother brought a wrongful death action against a manufacturer of the handgun used in the killing, seeking to recover on strict liability theories of an unreasonably dangerous product and ultrahazardous activity. In refusing to grant summary judgment for the manufacturer on the



ultrahazardous activity claim, the court said:

The defendant maintains that, if liability is imposed in this case, no company that markets handguns for sale to the general public will be able in the future to obtain insurance. The result, according to the defendant, will be catastrophic for handgun manufacturers: all such companies will be forced either to alter their marketing practices radically or to go out of business. This argument has a ring of plausibility to it. At the same time, however, it is highly speculative.

\* \* \*

Perhaps the most significant fact the defendant ignores is that increased insurance costs can be passed on to consumers in the form of higher prices for handguns. The people who benefit most from marketing practices like the defendant's are handgun manufacturers and handgun purchasers. Innocent victims rarely, if ever, are beneficiaries. Consequently, it hardly seems unfair to require manufacturers and purchasers, rather than innocent victims, to pay for the risks those practices entail. Furthermore, economic efficiency seems to require the same result. In an important article on ultrahazardous activities and risk allocation, Professor Clarence Morris makes just this point. Morris, "Hazardous Enterprises and Risk Bearing Capacity," 61 Yale L.J. 1172 (1952). In his view, "the avowed goal of the absolute liability approach is allocation of loss to the party better equipped to pass it on to the public: the superior risk bearer." *Id.* at 1176. Professor Morris discusses a variety of examples to show that the defendant is not always the superior risk bearer in an ultrahazardous activity case. Here is what he says, however, about bodily injury and risk-bearing capacity:

The financial burden of disabling personal injury overwhelms most people. While many can bear the cost of minor injury, prolonged infirmity and extended medical expense often exceed the financial competence of common men. Unless [common man] happens to be rich or covered by one of the more generous workman's compensations plans, he will probably bear the risk less easily than Enterpriser. The preponderant likelihood is that Enterpriser is the better risk bearer of the two. *Id.* at 1177.

. . . Thus, both fairness and economic efficiency suggest that the community would be better off if the defendant's marketing practices were classified as ultrahazardous.

*Id.* at 202–204. How does this analysis square with that of Professor Priest? The Fifth Circuit reversed the district court's decision on appeal, holding that "[t]he marketing of handguns to the general public falls far beyond the boundaries of the Louisiana doctrine of ultrahazardous activities." *Perkins v. F.I.E. Corp.*, 762 F.2d 1250, 1268 (5th Cir. 1985). The court explained that "a ruling that the marketing of handguns constitutes an ultrahazardous activity 'would in practice drive manufacturers out of business' and 'would produce a handgun ban by judicial fiat.'" *Id.* at 1268–69.

**5. Other Goals of Tort Law: Appeasement:** Appeasement as a goal of tort law means that the purpose of the law is to limit the negative impact of the infliction of injury to the event of the injury itself. Tort law provides a way to right the wrong without the injured party retaliating through some destructive means. That is, the victim's vengeance is bought off by imposing tort liability on the wrongdoer. The victim is appeased in two ways: receipt of compensation and knowledge of the fact that the transgressor is punished by being required to pay.

**6. Other Goals of Tort Law: Justice and Liability:** The law of torts is sometimes viewed as the expression of a moral principle — one who by his fault has caused damage to another

ought to make compensation as a matter of justice. There are two views in support of this position — and either variant is simply a different way of saying the same thing. First, the principle of ethical retribution places emphasis upon the fact that the payment of compensation is harmful to the offender, and that justice requires that the offender suffer the harm. Second, the principle of ethical compensation looks at the same situation from the point of view of the victim. It emphasizes the fact that the payment of compensation is a benefit to the victim of the wrong, and declares that justice requires that the victim should receive this compensation. Regardless of the perspective one chooses, the policy implications are the same.

### **3. Allocative Efficiency**

The risk of injury associated with products or services is sometimes characterized as an externality imposed on third parties. In this view, tort law increases allocative efficiency by forcing the internalization of externalities. However, to the extent the risk of injury is known by the injured party, the market price should reflect the risk so that there is no externality. Nevertheless, under some circumstances, there might be questions about the ability of the market price to reflect the risk.

#### **Doe v. Miles Laboratories, Inc.**

United States District Court for the District of Maryland  
675 F. Supp. 1466 (1987)

RAMSEY, District Judge.

A plague inflicts society and this Court is called upon to adjudicate the extent to which the effects will be visited upon its victims. The facts are tragic. In the autumn of 1983, plaintiff Jane Doe, who a week previous had given birth, sought emergency medical treatment for vaginal bleeding. During the course of treatment, the attending physician ordered the administration of 500 units of "Konyne," a blood-coagulation-factor concentrate produced by Cutter Laboratories, a division of Miles. Treatment appeared successful and plaintiff eventually was discharged.

Over the course of the months to follow, plaintiff suffered from a succession of ailments, ultimately being diagnosed as infected by the HTLV-III virus, and as having Acquired Immuno-Deficiency Syndrome-Related Complex (ARC), a predecessor of AIDS. On July 6, 1986, plaintiffs Jane and John Doe filed suit, alleging claims for strict liability in tort, for breach of warranties, and for loss of consortium. Later plaintiffs amended the complaint to include negligence counts, and for punitive damages. Defendant Miles, following other procedural actions, filed this motion for summary judgment on plaintiffs' counts for breach of warranties, for strict liability in tort, and for strict liability in tort — failure to warn; and further seeks summary judgment on the counts for loss of consortium and punitive damages to the extent they are derivative of the first three.

\* \* \*

#### **Products Liability Law**

Defective products cause accidents that result in both economic losses and injuries either to persons or property. Allowing victims to recover for such losses was long a controversial issue. Indeed, the common law has followed a confusing and torturous path in perceiving and remedying the situation.

\* \* \*

Once liability in negligence became established, the concept of strict products liability gained favor as an alternative theory of recovery for injuries from defective products. It is

commonly stated that there are three reasons for holding manufacturers and dealers strictly liable for personal or property injury caused by defective products. First, innocent victims should not be forced to bear the costs of accidents, which still occurs far too often, for even a negligence action may impose an evidentiary burden impossible to meet. Second, that strict liability promotes accident prevention, for the manufacturers are in a better position to ascertain and control the risks associated with their products. Third, that manufacturers are in a better position than victims to bear the costs, for they can distribute the losses across the many who purchase the product, whereas an individual victim, unless he or she is exceptionally well-to-do or heavily insured, will be driven into bankruptcy or into social welfare programs.

Implicit in the above justification for strict products liability, though perhaps not clearly articulated, is a fourth argument, namely that strict products liability can promote the efficient allocation of resources. Society has chosen to allow market forces to set the price for goods and thus to determine their availability and distribution. In some respects the market is very efficient. The price purchasers pay invariably reflects direct costs such as raw products, capital investment, labor, plus a reasonable rate of return. However, in other respects the market is not efficient. Prices often do not reflect indirect costs. These hidden costs can include the effects of pollution or the expenses of accidents, and are what economists refer to as "externalities."

When the price of an item does not reflect both its direct costs and its externalities, the price will be lower than its actual cost. This lower price will stimulate an inefficient allocation of resources, for persons will be encouraged to buy more of the product than they might if they were paying its true price. Society thus may increase the consumption of the very goods that create pollution, and thus have indirect cleanup costs, or that are defective, and thus have indirect accident costs. Strict products liability shifts the cost back to manufacturers, who will then reprice the goods to reflect their actual costs. Strict products liability therefore affords society a mechanism for a rational allocation of resources.<sup>5</sup> Absent it, the costs of externalities are thrust upon victims or upon society through its governmental welfare programs. In essence, without it there is a subsidy given to the polluting or defective product.

\* \* \*

Whatever the theory of recovery, whether negligence or strict liability, it is now clear that the test in products liability is the same. A plaintiff must show 1) the existence of a defect; 2) the attribution of the defect to the seller; and 3) a causal relation between the defect and the injury.

\* \* \*

#### STRICT LIABILITY IN TORT

Defendant argues . . . for an exemption from strict liability in tort for blood or blood products. . . .

\* \* \*

Do policy considerations warrant exempting blood and blood products from strict

---

<sup>5</sup> The argument is often made that strict products liability has the potential to bankrupt manufacturers. Such an argument misses the salutary economic role strict products liability plays. Understood properly, it can be seen that strict liability promotes a rational market place. Society cannot make rational decisions concerning the allocation of resources unless the price reflects the true costs. When the price rises greatly, reflecting the fact the product produces either substantial direct costs or creates widespread externalities, it is rational to discourage or even abandon consumption of that product. Strict products liability thus allows the marketplace to make better informed decisions.

products liability in tort? Defendant argues that the "unavoidably unsafe products" exemption provided to § 402A by Comment *k* applies to it. Comment *k* reads:

There are some products which, in the present state of human knowledge, are quite incapable of being made safe for their intended and ordinary use. These are especially common in the field of drugs. An example is the vaccine for the Pasteur treatment of rabies, which not uncommonly leads to very serious and even permanently injurious consequences when it is injected. Since the disease itself invariably results in a dreadful death, both the marketing and the use of the vaccine are fully justified, notwithstanding the unavoidably high degree of risk which they involve. Such a product, properly prepared and accompanied by proper directions and warnings, is not defective, nor is it *unreasonably* dangerous. The same is true of many other vaccines, drugs and the like, many of which for that very reason cannot legally be sold except to physicians, or under the prescription of a physician. . . . The seller of such products, again with the qualification that they are properly prepared and marketed, and proper warning is given where the situation calls for it, is not to be held to strict liability for unfortunate consequences attending to their use, merely because he has undertaken to supply the public with an apparently useful and desirable product, attended with a known but apparently reasonable danger.

Maryland courts have never expressly adopted Comment *k*. Several decisions in this federal district court, though, have relied on Comment *k*, holding that Maryland courts would adopt it if an appropriate case were before them. Those cases, however, involved prescription medications and did not address whether blood, especially blood infected with disease, fell within Comment *k*'s exemption.

This Court is not prepared to find that HTLV-III carrying blood presents a "reasonable danger" as Comment *k* requires. It is estimated that up to 95% of severe hemophiliacs test positive for exposure to the HTLV-III virus. The nearly complete exposure by the group most in need of coagulant-factors and the inevitably fatal nature of the disease for those who actually develop it are stark facts. The fact the virus was undetectable prior to 1985 is not a mitigating factor. The best view is to consider blood containing undetectable diseases to be a defective product and therefore that strict liability is applicable.

It is argued that providers of blood and blood products are promoting the general welfare by making possible improved health. It is argued that it is a fundamental social policy of the State of Maryland to promote the supply of blood and blood products. And it is argued that to allow strict products liability, which given the wide exposure to AIDS due to transfusions could create potentially substantial liability, would so raise costs of production that the supply of blood could be choked off.

The arguments are unpersuasive. . . . Those who choose to operate in the economic marketplace play by the rules applicable to all.

The arguments in favor of strict products liability apply as persuasively to blood and blood products as they do to any other product. First, there is no reason why victims of defective blood should bear the costs where victims of other defective products do not. Second, strict liability would provide the incentive to promote all possible accident prevention, for it is a rational business decision to keep costs down. Third, the producers are in a better position to spread the costs than are individual consumers. Finally, it makes for a more efficient allocation of social resources when the price of a transfusion of blood or blood products reflects its true costs.

Entrepreneurs by their nature are risk taking individuals. To the extent they need an incentive to engage in socially beneficial activities, the law already provides it in the form of a corporate shield on personal liability. To do as defendant argues, and exempt blood from strict liability would be to subsidize the product by forcing either victims or government through its social welfare programs to bear accident costs. In the absence of a clear expression on the part of the legislature of an intent to subsidize a particular product, it is not this Court's role to create the subsidy indirectly by carving out a Judge made exemption to strict products liability.

Accordingly, the Court will deny defendant's motion for summary judgment on plaintiffs' claim for strict products liability.

\* \* \*

### Notes and Questions

**1. *Strict Liability and Externalities:*** Markets characterized by zero transaction costs, perfect information, and risk-neutral participants are able to obtain the optimal level of accident prevention even in the absence of any liability standard. If information regarding risk is not perfectly communicated, then there is an economic justification for a strict liability standard. In general, however, a strict liability standard does not result in a greater level of accident prevention relative to a negligence standard. Thus, strict liability serves only to allocate responsibility for insuring against non-negligent accidents. Certainly, the presence of non-negligent accidents represents a cost that society would like to avoid. Economic efficiency suggests that such costs be borne by the party who can most cost effectively insure against non-negligent accidents. The court clearly assumes that Miles Laboratories can most cost effectively insure against such risks. Is there any justification for this conclusion? If someone other than Miles could provide the most cost-effective insurance, is it fair to say that such defendants receive a "subsidy" from the victims or government?

**2. *Externalities, Property Rights, and Coase Theorem:*** In general, externalities arise because of poorly defined or unenforceable property rights. What property rights were in question in *Miles Laboratories*? Did the supposed externalities arise because of poorly defined property rights, an inability to enforce such rights, or some other reason? What implications does the Coase Theorem have for this problem? Does strict liability allow the Coase Theorem to operate? Do you agree with the court's belief that its interpretation of strict liability supports a rational marketplace? In a sense, doesn't the court's externality approach really just obscure much of the efficiency analysis of alternative liability rules?

### 4. Tort Damages and Incentives

Losing tort defendants are generally subject to the rule that they must make their victims whole by paying sufficient damages to put plaintiffs in the same positions as before the tortious act. The possibility of being ordered to pay damages provides an economic incentive that alters the behavior of potential tortfeasors. Damages affect the PL portion of the Hand Formula, thereby shifting  $C^*$  and the optimal amount of prevention.

#### *a. Compensatory Damages*

Most of the damages awarded in individual tort cases are intended to make the plaintiff whole again, at least financially. Such **compensatory damages** usually consist of three major types of loss: (1) past and future medical expenses, (2) past and future economic loss, and (3) past and future pain and suffering.

### ***b. Punitive Damages***

**Punitive damages** are damages designed to punish individual defendants and to deter potential tortfeasors. The typical cases in which punitive damages could be awarded are intentional torts and negligence cases where the defendant's conduct fell to a level of "gross negligence" or "willful and wanton" disregard for the plaintiff's safety. The deterrence model of tort law suggests that punitive damages should also be awarded in cases where claims are unlikely to be brought against a culpable defendant (or damages are difficult to measure). That is, the basic economic justification for punitive damages is in offsetting the low probability of ultimate liability (due to low probabilities of detection, enforcement, or liability). *See* A. Mitchell Polinsky & Steven Shavell, *Punitive Damages: An Economic Analysis*, 111 *Harvard Law Review* 869, 890 (1998).

The deterrence model considers the potential tortfeasor's expected marginal benefits and marginal costs of the contemplated activity, where the expected costs are the probability of harm multiplied by the likely amount of harm ( $P \times L$ ). However, a complete analysis requires consideration of the likelihood that the injured parties will be able to recover — that is, the expected cost to the potential tortfeasor must be discounted by the likelihood that the tortfeasor will be detected and successfully sued. In fraud cases, for example, defendants typically seek not only to conceal their wrongdoing, but also to make it difficult to determine who committed the wrongdoing, so that the victim does not even discover the fraud or who victimized them. Awarding compensatory damages only in the cases in which the fraud is discovered will result in under-deterrence of the defendant's wrongful conduct. Punitive damages could therefore be awarded in an amount equal to the damages caused by the defendant to victims who are unlikely to sue. A second economic case for punitive damages is when the conduct being sanctioned has no redeeming societal value — e.g., the utility a thug receives from beating people up. Punitive damages awards for intentional torts would fit into this category. In cases of intentional torts, like fraud or assault, there is often overlap between the deterrent and punitive goals of criminal punishment and punitive damages. *See* Jeffery W. Grass, *The Penal Dimensions of Punitive Damages*, 12 *Hastings Constitutional Law Quarterly* 241 (1985) (arguing that punitive damages serve the same purpose as criminal law, but fail to provide the constitutional protections provided in criminal courts). Whether this is a good thing or a bad thing, it does potentially relieve pressure on the criminal justice system.

### **Mathias v. Accor Economy Lodging, Inc.**

United States Court of Appeals for the Seventh Circuit

347 F.3d 672 (2003)

POSNER, *Circuit Judge*.

The plaintiffs brought this diversity suit governed by Illinois law against affiliated entities (which the parties treat as a single entity, as shall we) that own and operate the "Motel 6" chain of hotels and motels. One of these hotels (now a "Red Roof Inn," though still owned by the defendant) is in downtown Chicago. The plaintiffs, a brother and sister, were guests there and were bitten by bedbugs, which are making a comeback in the U.S. as a consequence of more conservative use of pesticides. The plaintiffs claim that in allowing guests to be attacked by bedbugs in a motel that charges upwards of \$100 a day for a room and would not like to be mistaken for a flophouse, the defendant was guilty of "willful and wanton conduct" and thus under Illinois law is liable for punitive as well as compensatory damages. The jury agreed and awarded each plaintiff \$186,000 in punitive damages though only \$5,000 in compensatory

damages. The defendant appeals, complaining primarily about the punitive-damages award. . . .

The defendant argues that at worst it is guilty of simple negligence, and if this is right the plaintiffs were not entitled by Illinois law to any award of punitive damages. It also complains that the award was excessive — indeed that any award in excess of \$20,000 to each plaintiff would deprive the defendant of its property without due process of law. The first complaint has no possible merit, as the evidence of gross negligence, indeed of recklessness in the strong sense of an unjustifiable failure to avoid a *known* risk was amply shown. In 1998, EcoLab, the extermination service that the motel used, discovered bedbugs in several rooms in the motel and recommended that it be hired to spray every room, for which it would charge the motel only \$500; the motel refused. The next year, bedbugs were again discovered in a room but EcoLab was asked to spray just that room. The motel tried to negotiate "a building sweep [by EcoLab] free of charge," but, not surprisingly, the negotiation failed. By the spring of 2000, the motel's manager "started noticing that there were refunds being given by my desk clerks and reports coming back from the guests that there were ticks in the rooms and bugs in the rooms that were biting." She looked in some of the rooms and discovered bedbugs. The defendant asks us to disregard her testimony as that of a disgruntled ex-employee, but of course her credibility was for the jury, not the defendant, to determine.

Further incidents of guests being bitten by insects and demanding and receiving refunds led the manager to recommend to her superior in the company that the motel be closed while every room was sprayed, but this was refused. This superior, a district manager, was a management-level employee of the defendant, and his knowledge of the risk and failure to take effective steps either to eliminate it or to warn the motel's guests are imputed to his employer for purposes of determining whether the employer should be liable for punitive damages. The employer's liability for compensatory damages is of course automatic on the basis of the principle of respondeat superior, since the district manager was acting within the scope of his employment.

The infestation continued and began to reach farcical proportions, as when a guest, after complaining of having been bitten repeatedly by insects while asleep in his room in the hotel was moved to another room only to discover insects there; and within 18 minutes of being moved to a third room he discovered insects in that room as well and had to be moved still again. (Odd that at that point he didn't flee the motel.) By July, the motel's management was acknowledging to EcoLab that there was a "major problem with bed bugs" and that all that was being done about it was "chasing them from room to room." Desk clerks were instructed to call the "bedbugs" "ticks," apparently on the theory that customers would be less alarmed, though in fact ticks are more dangerous than bedbugs because they spread Lyme Disease and Rocky Mountain Spotted Fever. Rooms that the motel had placed on "Do not rent, bugs in room" status nevertheless were rented.

It was in November that the plaintiffs checked into the motel. They were given Room 504, even though the motel had classified the room as "DO NOT RENT UNTIL TREATED," and it had not been treated. Indeed, that night 190 of the hotel's 191 rooms were occupied, even though a number of them had been placed on the same don't-rent status as Room 504. . . .

Although bedbug bites are not as serious as the bites of some other insects, they are painful and unsightly. Motel 6 could not have rented any rooms at the prices it charged had it informed guests that the risk of being bitten by bedbugs was appreciable. Its failure either to warn guests or to take effective measures to eliminate the bedbugs amounted to fraud and probably to battery as well. . . . There was, in short, sufficient evidence of "willful and wanton

conduct" within the meaning that the Illinois courts assign to the term to permit an award of punitive damages in this case.

But in what amount? In arguing that \$20,000 was the maximum amount of punitive damages that a jury could constitutionally have awarded each plaintiff, the defendant points to the U.S. Supreme Court's recent statement that "few awards [of punitive damages] exceeding a single-digit ratio between punitive and compensatory damages, to a significant degree, will satisfy due process." *State Farm Mutual Automobile Ins. Co. v. Campbell*, 538 U.S. 408 (2003). The Court went on to suggest that "four times the amount of compensatory damages might be close to the line of constitutional impropriety." *Id.* Hence the defendant's proposed ceiling in this case of \$20,000, four times the compensatory damages awarded to each plaintiff. The ratio of punitive to compensatory damages determined by the jury was, in contrast, 37.2 to 1.

The Supreme Court did not, however, lay down a 4-to-1 or single-digit-ratio rule — it said merely that "there is a presumption against an award that has a 145-to-1 ratio" — and it would be unreasonable to do so. We must consider why punitive damages are awarded and why the Court has decided that due process requires that such awards be limited. The second question is easier to answer than the first. The term "punitive damages" implies punishment, and a standard principle of penal theory is that "the punishment should fit the crime" in the sense of being proportional to the wrongfulness of the defendant's action, though the principle is modified when the probability of detection is very low (a familiar example is the heavy fines for littering) or the crime is potentially lucrative (as in the case of trafficking in illegal drugs). Hence, with these qualifications, which in fact will figure in our analysis of this case, punitive damages should be proportional to the wrongfulness of the defendant's actions.

Another penal precept is that a defendant should have reasonable notice of the sanction for unlawful acts, so that he can make a rational determination of how to act; and so there have to be reasonably clear standards for determining the amount of punitive damages for particular wrongs.

And a third precept, the core of the Aristotelian notion of corrective justice, and more broadly of the principle of the rule of law, is that sanctions should be based on the wrong done rather than on the status of the defendant; a person is punished for what he does, not for who he is, even if the who is a huge corporation.

What follows from these principles, however, is that punitive damages should be admeasured by standards or rules rather than in a completely ad hoc manner, and this does not tell us what the maximum ratio of punitive to compensatory damages should be in a particular case. To determine that, we have to consider why punitive damages are awarded in the first place.

England's common law courts first confirmed their authority to award punitive damages in the eighteenth century, at a time when the institutional structure of criminal law enforcement was primitive and it made sense to leave certain minor crimes to be dealt with by the civil law. And still today one function of punitive-damages awards is to relieve the pressures on an overloaded system of criminal justice by providing a civil alternative to criminal prosecution of minor crimes. An example is deliberately spitting in a person's face, a criminal assault but because minor readily deterrable by the levying of what amounts to a civil fine through a suit for damages for the tort of battery. Compensatory damages would not do the trick in such a case, and this for three reasons: because they are difficult to determine in the case of acts that inflict largely dignitary harms; because in the spitting case they would be too slight to give the victim an incentive to sue, and he might decide instead to respond with violence — and an age-old



purpose of the law of torts is to provide a substitute for violent retaliation against wrongful injury — and because to limit the plaintiff to compensatory damages would enable the defendant to commit the offensive act with impunity provided that he was willing to pay, and again there would be a danger that his act would incite a breach of the peace by his victim.

When punitive damages are sought for billion-dollar oil spills and other huge economic injuries, the considerations that we have just canvassed fade. As the Court emphasized in *Campbell*, the fact that the plaintiffs in that case had been awarded very substantial compensatory damages — \$1 million for a dispute over insurance coverage — greatly reduced the need for giving them a huge award of punitive damages (\$145 million) as well in order to provide an effective remedy. Our case is closer to the spitting case. The defendant's behavior was outrageous but the compensable harm done was slight and at the same time difficult to quantify because a large element of it was emotional. And the defendant may well have profited from its misconduct because by concealing the infestation it was able to keep renting rooms. Refunds were frequent but may have cost less than the cost of closing the hotel for a thorough fumigation. The hotel's attempt to pass off the bedbugs as ticks, which some guests might ignorantly have thought less unhealthful, may have postponed the instituting of litigation to rectify the hotel's misconduct. The award of punitive damages in this case thus serves the additional purpose of limiting the defendant's ability to profit from its fraud by escaping detection and (private) prosecution. If a tortfeasor is "caught" only half the time he commits torts, then when he is caught he should be punished twice as heavily in order to make up for the times he gets away.

Finally, if the total stakes in the case were capped at \$50,000 (2 x [\$ 5,000 + \$20,000]), the plaintiffs might well have had difficulty financing this lawsuit. It is here that the defendant's aggregate net worth of \$1.6 billion becomes relevant. A defendant's wealth is not a sufficient basis for awarding punitive damages. That would be discriminatory and would violate the rule of law, as we explained earlier, by making punishment depend on status rather than conduct. Where wealth in the sense of resources enters is in enabling the defendant to mount an extremely aggressive defense against suits such as this and by doing so to make litigating against it very costly, which in turn may make it difficult for the plaintiffs to find a lawyer willing to handle their case, involving as it does only modest stakes, for the usual 33–40 percent contingent fee.

In other words, the defendant is investing in developing a reputation intended to deter plaintiffs. It is difficult otherwise to explain the great stubbornness with which it has defended this case, making a host of frivolous evidentiary arguments despite the very modest stakes even when the punitive damages awarded by the jury are included.

As a detail (the parties having made nothing of the point), we note that "net worth" is not the correct measure of a corporation's resources. It is an accounting artifact that reflects the allocation of ownership between equity and debt claimants. A firm financed largely by equity investors has a large "net worth" (= the value of the equity claims), while the identical firm financed largely by debt may have only a small net worth because accountants treat debt as a liability.

All things considered, we cannot say that the award of punitive damages was excessive, albeit the precise number chosen by the jury was arbitrary. It is probably not a coincidence that  $\$5,000 + \$186,000 = \$191,000/191 = \$1,000$ : i.e., \$1,000 per room in the hotel. But as there are no punitive-damages guidelines, corresponding to the federal and state sentencing guidelines, it is inevitable that the specific amount of punitive damages awarded whether by a judge or by a jury will be arbitrary. (Which is perhaps why the plaintiffs' lawyer did not suggest a number to the jury.) The judicial function is to police a range, not a point.

But it would have been helpful had the parties presented evidence concerning the regulatory or criminal penalties to which the defendant exposed itself by deliberately exposing its customers to a substantial risk of being bitten by bedbugs. That is an inquiry recommended by the Supreme Court. See *State Farm Mutual Automobile Ins. Co. v. Campbell*, 123 S. Ct. at 1520, 1526. But we do not think its omission invalidates the award. We can take judicial notice that deliberate exposure of hotel guests to the health risks created by insect infestations exposes the hotel's owner to sanctions under Illinois and Chicago law that in the aggregate are comparable in severity to that of the punitive damage award in this case.

"A person who causes bodily harm to or endangers the bodily safety of an individual by any means, commits reckless conduct if he performs recklessly the acts which cause the harm or endanger safety, whether they otherwise are lawful or unlawful." 720 ILCS 5/12-5(a). This is a misdemeanor, punishable by up to a year's imprisonment or a fine of \$2,500, or both. 720 ILCS 5/12-5(b); 730 ILCS 5/5-8-3(a)(1), 5/5-9-1(a)(2). Of course a corporation cannot be sent to prison, and \$2,500 is obviously much less than the \$186,000 awarded to each plaintiff in this case as punitive damages. But this is just the beginning. For, what is much more important, a Chicago hotel that permits unsanitary conditions to exist is subject to revocation of its license, without which it cannot operate. Chi. Munic. Code §§ 4-4-280, 4-208-020, 050, 060, 110. We are sure that the defendant would prefer to pay the punitive damages assessed in this case than to lose its license.

AFFIRMED.

### Notes and Questions

**1. Aren't They Designed to Punish?:** Punitive damages are designed to punish the firm that commits a malicious tort. However, the punitive damages should "bear some reasonable proportion to the actual damages." Why? Doesn't this view of punitive damages skew the entire purpose behind them? That is, punitive damages are designed to punish the tortious firm rather than have *anything* to do with the victim. In essence, this view of the application of punitive damages gives potential tortfeasors the incentive to act grossly negligent or maliciously as long as the damages caused by the defective product are lower than the profit earned from the sale of the product. Is a better view the one taken by the jury? They essentially strip the firm of all profit created by this unsafe design. Is this the correct approach? What are the problems with this application of punitive damages?

**2. Punitive Damages and Optimal Levels of Deterrence:** Recall the discussion earlier of the Hand Formula and the economics of negligence versus strict liability standards. Both standards, when applied properly from an economic perspective, result in defendants (and plaintiffs) taking the optimal level of precautions — the level of precaution where the marginal cost of the precaution equals the marginal benefit. But what if a judge or jury can impose punitive damages in addition to compensatory damages? One would expect the defendant to take more precautions, even though the marginal benefit of those precautions to society might be less than the marginal cost. The availability of punitive damages thus may lead to over-deterrence — defendants taking wasteful precautions or ceasing production of a useful product altogether. The concern with over-deterrence is even more pronounced when punitive damage awards are unpredictable; however, there is a great deal of scholarly debate about the unpredictability of punitive damages and its impact. See, e.g., Neil Vidmar and Mirya Holman, *The Frequency, Predictability, and Proportionality of Jury Awards of Punitive Damages in State Courts in 2005: A New Audit*, 43 *Suffolk University Law Review* 855 (2010); Theodore Eisenberg et al., *Juries*,

*Judges and Punitive Damages: An Empirical Study*, 87 Cornell Law Review 743 (2002); A. Mitchell Polinsky, *Are Punitive Damages Really Insignificant, Predictable, and Rational? A Comment on Eisenberg et al.*, 26 Journal of Legal Studies 663 (1997).

**3. *Intentional Torts, Unintentional Torts, and Statistics:*** An intentional tort involves a deliberate action that results in an injury. For example, a company commits an intentional tort when it knowingly makes false statements about the quality of a competitor's products. A negligent tort is an unintentional tort that arises from the failure to use reasonable care toward one to whom a duty is owed, which results in injury. Unintentional torts occur in a variety of business settings — ranging from slip-and-fall accidents in a showroom, to defectively designed products. For many types of business behavior, distinguishing between intentional and unintentional torts is not obvious. For example, a hand tool manufacturer that sells millions of power saws per year knows for a statistical fact — because of the large numbers involved — that a certain number of consumers per year will be injured by the saws. But such injuries are typically analyzed as unintentional torts — the manufacturer did not intend to injure any particular individual. Therefore, the distinction appears to be based, in part, on whether the wrongdoer knows the identity of the injured party prior to the occurrence of the tort. From an economic perspective, however, the cases plainly are distinguishable: in the false statement case, we are far less worried about over-deterrence than we are in the hand tool case. Unlike the making of false statements, the sale of hand tools, even when they are statistically certain to cause injury, has positive social value.

**4. *Juries and the Hand Formula:*** The Hand Formula not only provides a framework for determining liability, it also provides guidance to individuals and businesses about what risk level is acceptable to society. In essence, the Hand Formula suggests that businesses should go through the type of cost-benefit analysis that the tool manufacturer could have applied in the preceding note — for example, the tool manufacturer could have compared the cost of making the product safer with the expected value of injuries avoided. This may sound good in theory. And, without a doubt, businesses routinely engage in such calculations — indeed, as a society we want these tradeoffs to be considered. However, juries are not very understanding of such cold, hard calculations, and they tend to punish business for making the tradeoffs. Experimental evidence indicates that juries punish defendants who have carefully weighed costs and benefits. See Cass R. Sunstein et al., *Punitive Damages: How Juries Decide* (2002), finding that mock juries tended to agree in their moral judgments about the defendant's conduct, yet rendered erratic and unpredictable dollar awards. The study found that the process of jury deliberation produced a striking "severity shift" toward ever-higher awards, and that jurors tended to ignore instructions from the judges; showed "hindsight bias," believing that what happened should have been foreseen; and penalized corporations that had based their decisions on careful cost-benefit analyses. Although judges made many of the same errors, they performed better in some areas, suggesting that judges (or other specialists) may be better equipped than juries to decide punitive damages.

**5. *Tort Reform, Caps on Punitive Damages, and Insurance:*** The tort reform movement has championed several reforms in recent years. Some type of cap on punitive damages is almost always part of legislative reform packages. Is a cap on punitive damages a "silver bullet" to solve problems with products liability and tort litigation? Professor George Priest suggests that punitive damages are just part of the problem and, thus, caps on punitive damages cannot solve the problem:

. . . [V]arious tort reform statutes have been supported by a coalition of business

and insurance interests, chiefly on the simple ground that modern tort liability is excessive and unfair. Observers have not generally appreciated, however, that each of these reform provisions will affect insurance markets in a similar way: they reduce the variance in insurance risk pools. Obviously, caps on non-economic and punitive damages reduce the range of potential liability outcomes. Similarly, abrogation of the doctrine of joint and several liability in favor of strict comparative fault reduces the risk that any one of a group of joint defendants will ultimately be required to satisfy the entire judgment. Deducting first-party insurance benefits from tort judgments will also reduce risk pool variance. More generally, of course, altering liability standards to make recovery more difficult for plaintiffs will diminish expected liability. To the extent that variance in risk pools is reduced, third-party tort law insurance becomes more supportable.

These reforms, while helpful, constitute only partial contributions toward solution of the problems caused by modern tort law. The provision of insurance through tort law has undermined insurance markets. In my view, these markets will not be fully restored until these insurance issues are dealt with more systematically. The insurance function must be excised from tort law altogether. None of the recent statutory reforms achieves that effect.

\* \* \*

. . . [L]imitations on punitive damages may reduce risk pool variance to some extent. In my view, however, it is appropriate not only to restrain, but to prohibit, punitive damage awards in product liability and other tort contexts. Punitive damage awards can be justified only where there is some likelihood (1) that normal damage measures cannot measure loss accurately — such as in defamation cases; (2) where there is substantial difficulty in detecting the existence of the injury — such as in fraud or, perhaps, some antitrust actions; or (3) where other incentives are needed to stimulate litigation. For cases in which manufacturers or other providers have deliberately misrepresented product safety or effectiveness . . . punitive damages should be awarded on grounds of the fraudulent behavior itself, not on grounds of the defective character of the product. There is no further need to award punitive damages in typical products and service liability contexts. In this respect, a legislative maximum on punitive damage awards or limitations on the conditions under which punitive awards may be made, constitutes only a partial solution.

The source of the insurability crisis is not the level of damages alone. Rather, the diffuse and indiscriminate expansion of substantive tort liability has led to the unraveling of insurance markets in an increasing number of contexts. This unraveling can be arrested only if substantive standards of liability are redefined to focus exclusively on the accident reduction goal. In my view, modern tort law provides inadequate controls on the accident rate and simultaneously creates a tort law insurance regime that disrupts insurance markets and harms the poor. The objectives of tort law reform are uncontroversial: to reduce the accident rate and to provide a more coherent and comprehensive regime of compensation insurance. These objectives cannot be achieved by tinkering with damage measures and by limited changes in liability standards for particularly sympathetic sets of defendants, such as governmental entities, dramshops or non-profit organizations. Instead, modern tort law must be reformed systematically: by a complete redefinition of liability standards to better achieve accident reduction and insurance.

George Priest, *The Current Insurance Crisis and Modern Tort Law*, 96 Yale Law Journal 1521,

1587–90 (1987).

*c. The Collateral Source Rule*

The collateral source rule applies to situations where the victim receives compensation for his damages from a source independent of the tortfeasor. Under this rule, the payments received from the independent source are not deducted from the award the victim would otherwise receive from the tortfeasor. Therefore, the tortfeasor is not able to benefit from the victim's foresight in purchasing insurance. From an economic perspective, the collateral source rule raises two primary issues. First, what impact does the collateral source rule have on deterrence? Second, does the collateral source rule allow the victim a double recovery?

**Helfend v. Southern California Rapid Transit District**  
Supreme Court of California  
465 P.2d 61 (1970)

TOBRINER, Acting Chief Justice

\* \* \*

1. The facts.

Shortly before noon on July 19, 1965, plaintiff drove his car in central Los Angeles east on Third Street approaching Grandview. At this point Third Street has six lanes, four for traffic and one parking lane on each side of the thoroughfare. While traveling in the second lane from the curb, plaintiff observed an automobile driven by Glen A. Raney, Jr., stopping in his lane and preparing to back into a parking space. Plaintiff put out his left arm to signal the traffic behind him that he intended to stop; he then brought his vehicle to a halt so that the other driver could park.

At about this time Kenneth A. Mitchell, a bus driver for the Southern California Rapid Transit District, pulled out of a bus stop at the curb of Third Street and headed in the same direction as plaintiff. Approaching plaintiff's and Raney's cars which were stopped in the second lane from the curb, Mitchell pulled out into the lane closest to the center of the street in order to pass. The right rear of the bus sideswiped plaintiff's vehicle, knocking off the rearview mirror and crushing plaintiff's arm, which had been hanging down at the side of his car in the stopping signal position. . . . Plaintiff acquired some permanent discomfort but no permanent disability from the injuries sustained in the accident. . . .

Plaintiff filed a tort action against the Southern California Rapid Transit District, a public entity, and Mitchell, an employee of the transit district. At trial plaintiff claimed slightly more than \$2,700 in special damages, including \$921 in doctor's bills, a \$336.99 hospital bill, and about \$45 for medicines. Defendant requested permission to show that about 80 percent of the plaintiff's hospital bill had been paid by plaintiff's Blue Cross insurance carrier and that some of his other medical expenses may have been paid by other insurance. . . .

After the jury verdict in favor of plaintiff in the sum of \$16,300, defendants appealed, raising only two contentions: (1) The trial court committed prejudicial error in refusing to allow the introduction of evidence to the effect that a portion of the plaintiff's medical bills had been paid from a collateral source. (2) The trial court erred in denying defendant the opportunity to determine if plaintiff had been compensated from more than one collateral source for damages sustained in the accident.

We must decide whether the collateral source rule applies to tort actions involving public entities and public employees in which the plaintiff has received benefits from his medical

insurance coverage.

## 2. The collateral source rule.

The Supreme Court of California has long adhered to the doctrine that if an injured party receives some compensation for his injuries from a source wholly independent of the tortfeasor, such payment should not be deducted from the damages which the plaintiff would otherwise collect from the tortfeasor. As recently as August 1968 we unanimously reaffirmed our adherence to this doctrine, which is known as the "collateral source rule."

Although the collateral source rule remains generally accepted in the United States, nevertheless many other jurisdictions have restricted or repealed it. In this country most commentators have criticized the rule and called for its early demise. . . .

The collateral source rule as applied here embodies the venerable concept that a person who has invested years of insurance premiums to assure his medical care should receive the benefits of his thrift. The tortfeasor should not garner the benefits of his victim's providence.

The collateral source rule expresses a policy judgment in favor of encouraging citizens to purchase and maintain insurance for personal injuries and for other eventualities. Courts consider insurance a form of investment, the benefits of which become payable without respect to any other possible source of funds. If we were to permit a tortfeasor to mitigate damages with payments from plaintiff's insurance, plaintiff would be in a position inferior to that of having bought no insurance, because his payment of premiums would have earned no benefit. Defendant should not be able to avoid payment of full compensation for the injury inflicted merely because the victim has had the foresight to provide himself with insurance.

Some commentators object that the above approach to the collateral source rule provides plaintiff with a "double recovery," rewards him for the injury, and defeats the principle that damages should compensate the victim but not punish the tortfeasor. We agree with Professor Fleming's observation, however, that "double recovery is justified only in the face of some exceptional, supervening reason, as in the case of accident or life insurance, where it is felt unjust that the tortfeasor should take advantage of the thrift and prescience of the victim in having paid the premiums." (Fleming, *Introduction to the Law of Torts* (1967) p. 131.) . . .

Furthermore, insurance policies increasingly provide for either subrogation or refund or benefits upon a tort recovery, and such refund is indeed called for in the present case. (See Fleming, *The Collateral Source Rule and Loss Allocation in Tort Law*, *supra*, 54 Cal.L.Rev. 1478, 1479.) Hence, the plaintiff receives no double recovery; the collateral source rule simply serves as a means of by-passing the antiquated doctrine of non-assignment of tortious actions and permits a proper transfer of risk from the plaintiff's insurer to the tortfeasor by way of the victim's tort recovery. The double shift from the tortfeasor to the victim and then from the victim to his insurance carrier can normally occur with little cost in that the insurance carrier is often intimately involved in the initial litigation and quite automatically receives its part of the tort settlement or verdict.

Even in case in which the contract or the law precludes subrogation or refund of benefits, or in situations in which the collateral source waives such subrogation or refund, the rule performs entirely necessary functions in the computation of damages. For example, the cost of medical care often provides both attorneys and juries in tort cases with an important measure for assessing the plaintiff's general damages. To permit the defendant to tell the jury that the plaintiff has been recompensed by a collateral source for his medical costs might irretrievably upset the complex, delicate, and somewhat indefinable calculations which result in the normal jury verdict.

We also note that generally the jury is not informed that plaintiff's attorney will receive a

large portion of the plaintiff's recovery in contingent fees or that personal injury damages are not taxable to the plaintiff and are normally deductible by the defendant. Hence, the plaintiff rarely actually receives full compensation for his injuries as computed by the jury. The collateral source rule partially serves to compensate for the attorney's share and does not actually render 'double recovery' for the plaintiff. Indeed, many jurisdictions that have abolished or limited the collateral source rule have also established a means for assessing the plaintiff's costs for counsel directly against the defendant rather than imposing the contingent fee system. In sum, the plaintiff's recovery for his medical expenses from both the tortfeasor and his medical insurance program will not usually give him "double recovery," but partially provides a somewhat closer approximation to full compensation for his injuries.

If we consider the collateral source rule as applied here in the context of the entire American approach to the law of torts and damages, we find that the rule presently performs a number of legitimate and even indispensable functions. Without a thorough revolution in the American approach to torts and the consequent damages, the rule at least with respect to medical insurance benefits has become so integrated within our present system that its precipitous judicial nullification would work hardship. In this case the collateral source rule lies between two systems for the compensation of accident victims: the traditional tort recovery based on fault and the increasingly prevalent coverage based on non-fault insurance. Neither system possesses such universality of coverage or completeness of compensation that we can easily dispense with the collateral source rule's approach to meshing the two systems. The reforms which many academicians propose cannot easily be achieved through piecemeal common law development; the proposed changes, if desirable, would be more effectively accomplished through legislative reform. In any case, we cannot believe that the judicial repeal of the collateral source rule, as applied in the present case, would be the place to begin the needed changes.

\* \* \*

### **Notes and Questions**

**1. *The Collateral Source Rule and Deterrence:*** From a deterrence perspective, the collateral source rule reaches the right conclusion. Consider a defendant's incentives if allowed to set up the plaintiff's insurance coverage as a defense to his own responsibility in damages. Suppose that the expected harm from the accident equals \$500 and that this harm can be avoided by the defendant at a cost of \$400. The defendant's failure to incur these costs results in a finding of negligence under the Hand Formula. However, if the plaintiff's receipt of \$500 from the insurance company relieves the defendant of his damages, he no longer has an incentive to avoid this liability. Thus, potential tortfeasors tend to underinvest in accident prevention in the absence of the collateral source rule. Does this analysis require the potential tortfeasors to know whether their potential victims have purchased insurance?

**2. *Insurance Rates, Accident Avoidance Costs, and the Collateral Source Rule:*** Some potential tortfeasors may make investments in accident avoidance because of concern that accidents will cause their insurance rates to increase even if the collateral source rule did not preclude payment of the claim. How does this point affect the analysis in note 1?

**3. *Insurance Markets, Risk Pools, and the Collateral Source Rule:*** Abolition of the collateral source rule could improve the functioning of insurance markets because, as suggested by Professor Priest: "Deducting first-party insurance benefits from tort judgments will also reduce risk pool variance." Evaluate this argument.

**4. *The Collateral Source Rule and Double Recovery:*** The collateral source rule is not a free lunch for insured victims because insurance rates will adjust to reflect whether the insurance

company must cover the victim's expenses. The premiums that an insured pays are a reflection of the rights that the coverage purchases. Consider two different insurance contracts. First, an insurance company could offer a contract that obligated it to pay damages upon the injury of the insured as a result of a third party's negligence without providing the insurer with rights to sue the negligent third party. In short, under this contract, the insurer is stuck with the loss. However, the insurer will charge the insured a higher premium for this contract. Second, consider a contract that contains a term that assigns the legal rights of the insured to the insurer once the insured has been paid under the policy. Under this contract, the insurer has greater protection against bearing the loss and as a result will charge the insured a smaller premium. From this analysis, it is easy to see that when a victim is able to recover from both the insurance company and the defendant, the victim has paid for this so-called "double recovery" through higher insurance premiums. Therefore, it does not represent a windfall to the victim. However, are all insured parties better off with or without the collateral source rule?

**5. Fortuitous Benefits and the Collateral Source Rule:** What if, through the negligence of someone else, you are forced to change your plans, and in doing so meet the love of your life? Presumably, the benefit of marriage and a blissful life with your newfound love outweigh the damages done to you by the tortfeasor. However, the collateral source rule does not contemplate fortuitous benefits any more than it does collateral benefits. *See* Michael I. Krauss & Jeremy Kidd, *Collateral Source and Tort's Soul*, 48 University of Louisville Law Review 1, 33-35 (2009). Should you still be able to recover damages even though, in hindsight, you are happy to have incurred the loss and would gladly do so again?

## **E. Risk Regulation**

The earlier discussion of compensating wage differentials demonstrated that the market determined level of workplace safety is below a no-risk level of safety. At some point, workers, consumers, and investors voluntarily accept risk. However, in an increasing number of situations, politicians and government regulators have decided that the risk associated with particular products, services, or activities is unacceptably high and thus should be subjected to regulation. A large portion of the federal regulatory structure deals with health and safety regulations. Examples of agencies responsible for such regulation include the Occupational Safety and Health Administration (OSHA), the Environmental Protection Agency (EPA), the Consumer Product Safety Commission (CPSC), and the National Highway Traffic Safety Administration (NHTSA). Over the last thirty years, government risk regulation has become a big business. The purpose of this section is to provide an economic perspective on risk regulation.

### **1. Cost-Benefit Analysis**

Many regulators and the legislative mandates from which they operate contemplate risk-free levels of safety. However, because risk reduction is costly and involves tradeoffs, risk-free standards are not feasible. The practical impact of regulators pursuing risk-free standards is that cost considerations are often ignored. Rational economic decision making requires that government risk regulations be subjected to some type of cost-benefit analysis. In considering additional levels of risk reduction, the relevant economic analysis occurs at the margin. The marginal cost of reducing the risk associated with various activities increases as more risk is eliminated. Conversely, the marginal benefits of reduced risk decline as greater amounts of risk are eliminated. The optimal risk level occurs at the point where the marginal benefits of risk reduction equal the marginal costs. Beyond this point, risk reduction can occur only if allocative



efficiency is sacrificed.

**UAW v. Occupational Safety & Health Administration**  
United States Court of Appeals for the District of Columbia Circuit  
938 F.2d 1310 (1991)

STEPHEN F. WILLIAMS, Circuit Judge:

Representatives of labor and industry challenge a regulation of the Occupational Safety and Health Administration, "Control of Hazardous Energy Sources (Lockout/Tagout)," 54 *Fed.Reg.* 36,644 (1989). The regulation deals not with the effects of such subtle phenomena as electrical energy fields but with those of ordinary industrial equipment that may suddenly move and cut or crush or otherwise injure a worker. "Lockout" and "tagout" are two procedures designed to reduce these injuries. Lockout is the placement of a lock on an "energy isolating device", such as a circuit breaker, so that equipment cannot start up until the lock is removed. Tagout is the similar placement of a plastic tag to alert employees that the tagged equipment "may not be operated" until the tag is removed. Although OSHA had previously issued specific standards governing especially dangerous equipment, the present rule extends lockout/tagout to virtually all equipment in almost all industries. It generally requires employers to use lockout procedures during servicing and maintenance, unless the employer can show that tagout will provide the same level of safety.

\* \* \*

. . . [T]he National Association of Manufacturers [contends] that Congress has given so little guidance for rules issued under § 6(b) [of the Occupational Safety and Health Act] but *not* covered by § 6(b)(5) that as to such rules the Act invalidly delegates legislative authority. Although we reject that claim, we find that the interpretation offered by the Secretary is, in light of nondelegation principles, so broad as to be unreasonable. We note, however, the existence of at least one interpretation that is reasonable and consistent with the nondelegation doctrine.

\* \* \*

I

[The court first rejected petitioner UAW's contention that § 6(b)(5) provided the statutory criteria for the lockout rule.]

II

The removal of § 6(b)(5) as a direct constraint on OSHA regulation outside the area of toxics gives point to the NAM's claim of an excessive delegation of legislative power. The only evident source of constraints remaining is § 3(8). It defines an "occupational safety and health standard" as a standard which requires conditions, or the adoption or use of one or more practices, means, methods, operations, or processes, reasonably necessary or appropriate to provide safe or healthful employment and places of employment." Though the language is exceedingly vague, the *Benzene* plurality found it the source of a threshold requirement of "significant risk," without which OSHA was not to act under § 6(b) at all. It justified this narrowing construction with the argument (among others) that otherwise "the statute would make such a 'sweeping delegation of legislative power' that it might be unconstitutional under the Court's reasoning in *A.L.A. Schechter Poultry Corp. v. United States*, 295 U.S. 495 (1935). . . .

The *Benzene* construction was, of course, a manifestation of the Court's current general practice of applying the nondelegation doctrine mainly in the form of "giving narrow constructions to statutory delegations that might otherwise be thought to be unconstitutional." . . . In effect we require a clear statement by Congress that it intended to test the constitutional

waters.

We thus turn to possible constructions.

A

One can imagine broader constructions than the one proposed by OSHA, but not easily. .

\* \* \*

It is true that price and wage controls blanketing the entire economy have been sustained under quite vague legislative directions. But in view of the inevitable tensions in such controls between such purposes as price stabilization on the one hand and the need for adjustments on ground of changes in cost and other market conditions on the other, an insistence on greater clarity from Congress would deny it any power to impose price controls at all. Not so here. Congress can readily articulate some principle by which the beneficent health and safety effects of workplace regulation are to be traded off against the adverse welfare effects. "Policy direction is all that was ever required, and policy direction is what is lacking in much contemporary legislation." John Hart Ely, *Democracy and Distrust* 133 (1980). OSHA's reading of the Act finds no such direction.

We note that OSHA's claimed discretion is procedurally confined. The agency sets "standards," which would normally apply across an industry, or to a category of machines, or to some other reasonably broad category. Thus, even under its view OSHA would normally not be free to single out the Jones Company for standards embodying strict feasibility while letting the Smith Company off on ones reflecting some different principle. But even the use of general standards leaves opportunities for dangerous favoritism. The cost of compliance with a standard will vary among firms in an industry, so the power to vary the stringency of the standard is the power to decide which firms will live and which will die. At the simplest level, for example, compliance may involve economies of scale, so that a tough standard will erase small, marginal firms and leave the field to a small group of larger ones. Compare Ann P. Bartel & Lacy Glenn Thomas, "Direct and Indirect Effects of Regulation: A New Look at OSHA's Impact," 28 *J.L. & Econ.* 1, 23–25 (1985).

OSHA's proposed analysis would give the executive branch untrammelled power to dictate the vitality and even survival of whatever segments of American business it might choose. Although in *Benzene* the Court focused perhaps more on the severity of the power claimed by OSHA than on its variability, the plurality's point is apt here: "In the absence of a clear mandate in the Act, it is unreasonable to assume that Congress intended to give the Secretary the unprecedented power over American industry that would result from the Government's view. . . ." 448 U.S. at 645. At least if reasonable alternative readings can be found, OSHA's must be rejected as unreasonable.

B

The NAM argues (as a fallback to its nondelegation claim) that Congress's use of "reasonably necessary or appropriate" in § 3(8) contemplates "cost-benefit" analysis. Under this interpretation, in imposing standards under § 6(b) but outside the realm of toxics, OSHA may adopt a safety standard if its benefits outweigh its costs, and not otherwise.

Cost-benefit analysis is certainly consistent with the language of § 3(8). "Reasonableness" has long been associated with the balancing of costs and benefits. The "reasonable" person of tort fame is one who takes a precaution if the gravity of the injuries averted, adjusted for their probability, exceeds the precaution's burden. *United States v. Carroll Towing Co.*, 159 F.2d 169, 173 (2d Cir.1947).

And while the legislative history is almost blank on the subject, it suggests concern with market failures, and properly conducted cost-benefit analysis should yield a solution approximating that of a market undistorted by market failures.<sup>6</sup> Application of cost-benefit analysis to safety standards also gives effect to Congress's distinction between slow-acting hazards and others, with its "particular concern for health hazards of 'unprecedented complexity' that had resulted from chemicals whose toxic effects 'are only now being discovered.'"

Moreover, courts have often taken the word "reasonable" in a statute to require that burdens be justified by the resulting benefits. . . .

\* \* \*

The union argues that prior cases preclude a cost-benefit interpretation here. They do not. The Supreme Court has expressly reserved the question. In *Cotton Dust*, upholding the Secretary's understanding that the "feasib[ility]" criterion of § 6(b)(5) was not a cost-benefit standard, the Court observed that "[w]hen Congress has intended that an agency engage in cost-benefit analysis, it has clearly indicated such intent on the face of the statute." 452 U.S. at 510. But it cited approvingly *Forester* and *Aqua Slide 'N' Dive* (which find "unreasonable risk" to incorporate cost-benefit balancing), *id.* at 510–11 n. 30, and went on explicitly to leave open the question of § 3(8)'s meaning apart from toxics regulation: "This is not to say that § 3(8) might not require the balancing of costs and benefits for standards promulgated under provisions other than § 6(b)(5) of the Act." *Id.* at 513 n. 32.

\* \* \*

As there appear to be many confusions about cost-benefit analysis, it may be important to make clear what we are not saying when we identify it as a reasonable interpretation of § 3(8) as applied outside the § 6(b)(5) realm. Cost-benefit analysis requires identifying values for lost years of human life and for suffering and other losses from non-fatal injuries. Nothing we say here should be taken as confining the discretion of OSHA to choose among reasonable evaluation methods. While critics of cost-benefit analysis argue that any such valuation is impossible, that is so only in the sense that pin-point figures are necessarily arbitrary, so that the decisionmaker is effectively limited to considering some range of values. In fact, we make implicit life and safety valuations each day when we decide, for example, whether to travel by train or car, the former being more costly (at least if several family members are traveling together) but safer per passenger-mile. Where government makes decisions for others, it may reasonably be expected to make the trade-offs somewhat more explicitly than individuals choosing for themselves. The difficulty of securing agreement even on a range of values hardly justifies making decisions on the basis of a pretense that resources are not scarce. In any event, OSHA has an existing obligation under Executive Order No. 12,291, 46 Fed.Reg. 13,193 (1981), to complete a cost-benefit analysis for each major rulemaking, so use of such a standard not only

---

<sup>6</sup> Indeed, the Regulatory Impact Analysis of the Lockout/Tagout rule assessed it precisely on that basis, identifying worker lack of information and immobility as the relevant sources of market failure. Compare generally W. Kip Viscusi, *Risk by Choice* (1983), and especially *id.* at 44 (finding the safety incentives created by observed job-risk premiums nearly 3000 times stronger than those created by OSHA fines), and *id.* at 77 (finding systemic preference for jobs whose risks are not readily understood). As OSHA never adopted the Regulatory Impact Analysis as its own reasoning, we cannot treat the rule as an application of cost-benefit analysis; accordingly we do not review the Regulatory Impact Analysis to determine whether it would satisfy that standard.

is doable in the qualified sense of which we have spoken but can be done without additional regulatory resources.

Thus, cost-benefit analysis entails only a systematic weighing of pros and cons, or what Benjamin Franklin referred to as a "moral or prudential algebra." Writing to a friend who was perplexed by a difficult decision, he explained his own approach:

When those difficult cases occur, they are difficult, chiefly because while we have them under consideration, all the reasons pro and con are not present to the mind at the same time. . . . To get over this, my way is to divide half a sheet of paper by a line into two columns; writing over the one Pro, and over the other Con. Then, during three or four days consideration, I put down under the different heads short hints of the different motives, that at different times occur to me, for or against the measure. When I have thus got them all together in one view, I endeavor to estimate their respective weights. . . . And, though the weight of reasons cannot be taken with the precision of algebraic quantities, yet when each is thus considered, separately and comparatively, and the whole lies before me, I think I can judge better, and am less liable to make a rash step, and in fact I have found great advantage from this kind of equation, in what may be called moral or prudential algebra.

Reprinted in Edward M. Gramlich, *Benefit-Cost Analysis of Government Programs* 1–2 (1981).

As we accept the NAM's contention that § 3(8)'s "reasonably necessary or appropriate" criterion can reasonably be read as requiring cost-benefit analysis, we must reject its nondelegation claim.

We hold only that cost-benefit is a permissible interpretation of § 3(8). Given the ambiguity inherent in that section, there may be other interpretations that conform to nondelegation principles. Accordingly we remand to OSHA, noting that its treatment of some of the parties' other claims, discussed below, will likely turn on its decision. We note, however, that Executive Order No. 12,291 may bear on OSHA's authority to promulgate a safety standard whose benefits fail to outweigh its costs. Section 2 of that order provides:

In promulgating new regulations . . . all agencies, *to the extent permitted by law, shall adhere to the following requirements. . . .*

(b) *Regulatory action shall not be undertaken unless the potential benefits to society for the regulation outweigh the potential costs to society;*

(c) *Regulatory objectives shall be chosen to maximize the net benefits to society. .*

46 Fed.Reg. 13,193 (1981) (emphasis added).

\* \* \*

Accordingly, in light of the NAM's nondelegation claim we reject OSHA's view that under § 3(8) it may impose any restriction it chooses so long as it is "feasible," but we also reject the NAM's nondelegation claim in light of our view that § 3(8) may reasonably be read as providing for cost-benefit analysis.

\* \* \*

Accordingly, we remand the case to OSHA for further consideration in light of this opinion.

So ordered.

STEPHEN F. WILLIAMS, Circuit Judge, concurring:

I write separately to address the UAW's apparent assumption that application of the significant risk/feasibility analysis associated with § 6(b)(5) is necessarily more protective of

health and safety than a cost-benefit criterion. This is not self-evidently true.

First, if OSHA applies cost-benefit analysis, then more risks seem likely to qualify as "significant" within the meaning of *Benzene*; many risks that may seem insignificant if their discovery triggers regulatory burdens limited only by feasibility, as under § 6(b)(5), may be significant if the consequence is cost-justified corrective measures.

Second, even where the application of cost-benefit analysis would result in less stringent regulation, the reduced stringency is not necessarily adverse to health or safety. More regulation means some combination of reduced value of firms, higher product prices, fewer jobs in the regulated industry, and lower cash wages. All the latter three stretch workers' budgets tighter (as does the first to the extent that the firms' stock is held in workers' pension trusts). And larger incomes enable people to lead safer lives. One study finds a 1 percent increase in income associated with a mortality reduction of about 0.05 percent. Jack Hadley & Anthony Osei, "Does Income Affect Mortality?," 20 *Medical Care* 901, 913 (September 1982). Another suggests that each \$7.5 million of costs generated by regulation may, under certain assumptions, induce one fatality. Ralph L. Keeney, "Mortality Risks Induced by Economic Expenditures," 10 *Risk Analysis* 147, 155 (1990) (relying on E.M. Kitagawa & P.M. Hauser, *Differential Mortality in the United States of America: A Study of Socioeconomic Epidemiology*(1973)). Larger incomes can produce health by enlarging a person's access to better diet, preventive medical care, safer cars, greater leisure, etc. See Aaron Wildavsky, *Searching for Safety* 59–71 (1988).

Of course, other causal relations may be at work too. Healthier people may be able to earn higher income, and characteristics and advantages that facilitate high earnings (e.g., work ethic, education) may also lead to better health. Compare C.P. Wen, et al., "Anatomy of the Healthy Worker Effect: A Critical Review," 25 *J. of Occupation Medicine* 283 (1983). Nonetheless, higher income can secure better health, and there is no basis for a casual assumption that more stringent regulation will always save lives.

It follows that while officials involved in health or safety regulation may naturally be hesitant to set any kind of numerical value on human life, undue squeamishness may be deadly. Incremental safety regulation reduces incomes and thus may exact a cost in human lives. For example, if analysis showed that "an individual life was lost for every \$12 million taken from individuals [as a result of the regulation], this would be a guide to a reasonable value tradeoff for many programs designed to save lives." Keeney, "Mortality Risks Induced by Economic Expenditures," 10 *Risk Analysis* at 158. Such a figure could serve as a ceiling for value-of-life calculated by other means, since regulation causing greater expenditures per life expected to be saved would, everything else being equal, result in a net loss of life.

### Notes and Questions

**1. Marginal Cost-Marginal Benefit Analysis:** Cost-benefit analysis, when performed properly, should be an analysis of marginal costs and marginal benefits. Reliance on average costs and average benefits can lead to mistakes. Consider, for example, the Environmental Protection Agency's regulation of acceptable levels of exposure to arsenic.<sup>7</sup> Table VI-1 shows the average and marginal costs per life saved by moving from relatively loose to strict standards. Assume the regulators are using \$7.5 million per life saved as the marginal benefit. (Note that average benefit is equal to marginal benefit because the marginal benefit is constant.).

---

<sup>7</sup> This example is from W. Kip Viscusi et al., *Economics of Regulation and Antitrust* 667–68 (2d ed. 1995).

Comparison of average cost per life saved to the value of the life saved indicates that the strict standard should be adopted. However, the comparison of marginal cost per life saved indicates that it would be very wasteful to require the strict standard. The loose standard would be optimal. If policy decisions were guided by the average cost per life saved (which was substantially below the standard value of life measurements used in risk management), an economically inefficient level of regulation would result.

[INSERT TABLE VI-1 ABOUT HERE]

**2. Risk Assessment and Risk Management:** When analyzing cases that involve cost-benefit analysis, it is helpful to set the framework for the analysis by considering the areas of risk assessment and risk management. These important parts of the regulatory system are described by Judge (now Justice) Stephen Breyer in his influential book, *Breaking the Vicious Cycle: Toward Effective Risk Regulation*:

The [regulatory] system has two basic parts, a technical part, called "risk assessment," designed to measure the risk associated with the substance, and a more policy-oriented part, called "risk management," which decides what to do about it.

Risk assessment can itself be divided into four activities: (a) *Identifying the potential hazard*, say benzene in respect to cancer: Is it benzene in any context, or just benzene used in industry, or undiluted benzene, or certain solutions of benzene in certain places? (b) *Drawing a dose/response curve*: How does the risk of harm vary with the person's exposure to that substance? The question is critically important, for, as Paracelsus pointed out over four hundred years ago, "the dose alone determines the poison." Drinking a bottle of pure iodine is deadly; placing a drop of diluted iodine on a cut is helpful. Regulators will try to use statistical studies of, say, cancer in humans (epidemiological studies), or experiments with high substance-doses given to animals, to estimate the potential effects of human exposure to low substance-doses over varying periods of time. (c) *Estimating the amount of human exposure*: How many persons in a particular workforce, or in a particular region, or in the public generally, will be exposed to different doses of the substance, and for how long? Suppose exposure to a solution of five parts per million every day for twenty years will be likely to cause five extra deaths per year per million persons. Exposing the entire population may mean 1,250 extra deaths; exposing a hundred thousand persons may mean one extra death every two years. And even in the latter case, if only two persons are exposed to the substance, so that each runs a 50 percent risk of death, there may be a regulatory problem. (d) *Categorizing the result*: Is the substance, in fact, a carcinogen? A strong carcinogen or a weak carcinogen? Based upon the dose/response and exposure findings, how should the risk assessment describe (or categorize) the hazard? In carrying out these activities, particularly in making dose/response and exposure estimates, regulators often find that they simply lack critically important scientific or empirical data: they do not know how many Americans inhale how much benzene at gasoline stations; they do not know the extent to which the biology of a rat or mouse resembles, or differs from, that of a human being. In such instances, they will often make a "default assumption" — a formalized guess — designed to fill the gap and to permit the regulator to continue the analysis.

*Risk management* determines what the regulator should do about the risks that the assessment reveals. Ideally, the risk manager will consider what will be likely in fact to occur should he choose each of several regulatory potions. On the one hand, to what extent will the regulation actually diminish the specific risk at issue? On the other hand,

to what extent will regulation itself produce *different* risks? (Will childproof aspirin bottle tops save children, or will they lead many parents, unable to open the top easily, simply to leave the top off? Will saccharin users, denied saccharin, switch to sugar, gain weight, and die of heart attacks?) To what extent will the regulation deprive users of benefits the substance now brings? To what extent will it impose added costs? The manager also must consider practical problems, such as the difficulty of enforcing a regulation or the political reaction that its promulgation might bring. Ultimately, in light of the identified risks, the risks associated with alternatives, the effect on benefits, the costs, and the practicalities, the risk managing regulator will reach a decision.

Stephen G. Breyer, *Breaking the Vicious Cycle: Toward Effective Risk Regulation* 9–10 (1993).

### **3. *Cost-Benefit Analysis, Compensating Wage Differentials, and Voluntary Risk***

**Taking:** A safety or health standard is socially desirable only if the value workers place on risk reduction is commensurate with the cost of complying with the standard. The value workers place on risk reduction can be estimated, in theory at least, from knowledge of compensating wage differentials. However, even if compensating wage differentials were accurately calculated from observations in a market characterized by both widespread information and choice, there are still objections to their use in assessing the benefits of OSHA standards. First, compensating wage differentials reflect the preferences of only those directly involved in the contractual relationship between the employer and employee. It is frequently argued that members of society who are not directly affected by the risk reduction program might be willing to pay something for the benefits that accrue to those who are directly affected. Presumably, this willingness to pay is strongest for family members, relatives, and close friends and weakest for strangers. However, even strangers would have some interest in reducing injury and disease if they were to be taxed in order to subsidize the medical treatment of those who are injured or become ill. Thus, it is argued the benefits of OSHA standards extend beyond the direct beneficiaries to other external parties whose willingness to pay should also be counted. The second argument against using only the apparent willingness of workers to pay for risk reduction as a measure of its benefit is that workers may not really know what's best for themselves in the long run. Society frequently prohibits, or at least tries to prohibit, people from indulging in activities that are dangerous to their welfare. Laws against the use of narcotics and gambling are two examples. Some argue that OSHA standards limiting exposure to dangerous substances or situations fall under the category of preventing workers from doing harm to themselves by being lured into dangerous work. Therefore, it is argued, to ask how much they value risk reduction is irrelevant. The conflict between those who claim that workers know what is best for themselves and those who claim they do not, can only be resolved on philosophical grounds. A final argument is that workers often lack negotiating power in the labor market. Say, for example, that you are an uneducated coal miner in a small mining community with high unemployment and few other economic opportunities. Your boss decides that she does not want to pay for gas masks anymore and that she is cutting your pay by five percent. Do you have any recourse?

### **4. *Worker Safety, Compensating Wage Differentials, and the Best Interests of***

**Workers:** In 1970, Congress passed the Occupational Safety and Health Act which directed the U.S. Department of Labor to issue and enforce safety and health standards for all private employers. The stated goal of the Act was to ensure the "highest degree of health and safety protection for the employee." Despite the ideal that employees should face the minimum possible risk in the work place, implementing this ideal as social policy is not necessarily in the best interest of workers. Some workers are more willing than others to perform risky jobs, and they

are usually rewarded with a compensating wage differential. When the government mandates the reduction of risk in a market where workers are compensated for the risk they take, it penalizes workers who are not terribly sensitive to risk and appreciate the higher wages associated with higher risk. The critical issue, of course, is whether workers have the knowledge and choice necessary to generate compensating wage differentials. Many people believe that workers are uninformed, unable to comprehend different risk levels, or immobile and thus that most do not choose risky jobs voluntarily. If this belief were true, government regulations could make workers better off. Indeed, while the evidence of a positive relationship between wage and risk of death should challenge the notion that information and mobility are generally insufficient to create compensating differentials, there are specific areas in which problems obviously exist. For example, the introduction each year of new workplace chemicals whose effects on humans may be unknown for two or more decades, owing to long gestation of most cancers and lung diseases, clearly presents substantial information problems to affected labor market participants. To say that worker utility can be increased by government regulation does not then imply that it will be increased. The outcome depends on how well the unregulated market functions and how careful the government is in setting its standards for risk reduction.

**5. OSHA's Feasibility Standard:** OSHA proposes the use of a feasibility test in setting regulatory standards. Regulatory stringency under such a test can be characterized as requiring only that technology-based standards be affordable (i.e. capable of being done). Clearly then, a feasibility test does not contemplate a balancing of the marginal costs and marginal benefits. Rather, under a feasibility standard, OSHA simply looks for a "kink" in the marginal cost curve. That is, at what point does it become prohibitively costly to raise the standard. Such a kink may of course occur past the point at which marginal benefits equal marginal costs. What justification can OSHA offer for a feasibility test? Does Judge Williams' concurring opinion adequately address such arguments? The Supreme Court's *Cotton Dust* opinion allows OSHA to use a feasibility standard in the regulation of toxics. What justification allows the feasibility standard to exist in the case of workplace toxics?

**6. The Impact of Risk Regulation:** Questions regarding the effectiveness of the current regulatory system don't stop after applying a cost-benefit test to stringency levels. In order to justify the existence of regulatory agencies, the regulations should have the desired impact on market behavior:

Firms will choose to make the necessary investments in health and safety if the OSHA enforcement policy in conjunction with market incentives for safety makes it in the firm's financial self-interest to do so. More specifically, a firm will comply with an OSHA regulation if

$$\text{Expected cost of compliance} < \text{Probability of inspection} \times \text{Expected no. of violations per inspection} \times \text{Average penalty per violation}$$

As discussed, the three links in establishing these incentives — inspections, violations, and penalties — are all relatively weak. A firm has less than one chance in 200 of being inspected in any given year. If inspected, it expects to be found guilty of less than two violations of the standards, and for each violation the average penalty is under \$60. Overall, the financial cost per worker is just over fifty cents. A useful comparison is that market forces through compensating differentials in combination with workers' compensation premiums imposed costs in excess of \$800 per worker for the same period. Quite simply, OSHA's enforcement effort is too modest to create truly effective financial



incentives for safety.

W. Kip Viscusi et al., *Economics of Regulation and Antitrust* 816 (2d ed. 1995). A simple way of stating this is that firms will not comply with regulations if the cost of compliance is lower than the expected cost of getting caught. Based on this information what justification is there for the continued existence of OSHA? Does the market have a comparative advantage in allocating the types of risk that OSHA is attempting to regulate? It appears that in regard to both stringency and enforcement, regulatory agencies like OSHA have ignored the reality of the market.

**7. *Wealthier is Healthier:*** Judge Stephen Williams, in his concurring opinion, cautions that the costs of health and safety regulations might actually result in more lives lost than saved. This counter-intuitive result flows from the empirical evidence that shows strong correlations between average national income levels and measures of national health, such as life expectancy and infant mortality. Wealthier is healthier. Moreover, regulations — even those that are supposed to reduce risk — that do not pass a cost-benefit analysis can reduce wealth and, thus, reduce health. Jobs are lost and wealth is destroyed when regulations force businesses to make inefficient expenditures. Similarly, new product innovation is suppressed whenever government overregulates the development, distribution, and use of products with life-saving potential. Risk experts Richard J. Zeckhauser and W. Kip Viscusi argue that "overreaction to very small risks impedes the kind of technological progress that has historically brought dramatic improvements in both health and material well-being." Richard J. Zeckhauser & W. Kip Viscusi, *Risk Within Reason*, 248 *Science* 559 (1990). Does the court give adequate attention to such facts? What incentives would drive OSHA to ignore such facts?

**8. *Marginal Analysis and Heterogeneity:*** Use of marginal analysis suggests something about the manner in which regulatory standards should be applied across different industries or firms. Consider the case of a regulation that would be applied across several different industries. According to the regulator, the marginal benefit per unit of safety is the same across all industries. However, it is unlikely that firms in each of these different industries face the same marginal cost curves. Therefore, across industries, the level of safety at which marginal benefits equal marginal costs will be different. Suppose that the regulating agency sets safety standards based on a representative firm. If a single standard is imposed regardless of industry, then the regulating agency loses the opportunity for further risk reduction. If a firm has lower marginal costs than contemplated by the regulation, then it would be efficient for this firm to be required to have a higher safety standard. Thus, a single standard misses inexpensive opportunities to reduce risk. On the other hand, if a firm has higher marginal costs than supposed by the regulation, then the single standard will result in economic inefficiency. This results in lower incomes and therefore sacrificed safety. By using marginal analysis, regulatory agencies would be able to maximize the risk-reducing benefits of their regulations across industries and across firms.

## **2. Measuring Risk: All Risks Are Relative**

Absolute levels of risk provide very little relevant information in the absence of some relative comparison. For example, suppose that living five years at the boundary of a nuclear power plant site increased one's annual death risk by one chance in one million. Should you be concerned? Should there be additional regulation? Now consider the fact that the risk of traveling six minutes by canoe will also increase your annual death risk one chance in one million. Both of these risks have the same impact on annual death rates. However, many individuals might be outraged if tax dollars were being spent to eliminate the risk of being killed during a six minute

canoe trip. By considering a risk that sounds intimidating relative to risks that individuals freely undertake daily, it becomes much easier to achieve a conceptual understanding of risk measures. Gaining a conceptual measure of risk is important because regulatory agencies often attempt to regulate risks that are very small relative to the types of risks that individuals expose themselves to in their everyday lives. Furthermore, gaining an appreciation for the relative impacts of different risk reduction proposals helps in analyzing the efficacy of regulatory performance.

**Corrosion Proof Fittings v. Environmental Protection Agency**  
United States Court of Appeals for the Fifth Circuit  
947 F.2d 1201 (1991)

JERRY E. SMITH, Circuit Judge:

The Environmental Protection Agency (EPA) issued a final rule under section 6 of the Toxic Substances Control Act (TSCA) to prohibit the future manufacture, importation, processing, and distribution of asbestos in almost all products. Petitioners claim that . . . the rule was not promulgated on the basis of substantial evidence. . . . Because the EPA failed to muster substantial evidence to support its rule, we remand this matter to the EPA for further consideration in light of this opinion.

I. Facts and Procedural History.

Asbestos is a naturally occurring fibrous material that resists fire and most solvents. Its major uses include heat-resistant insulators, cements, building materials, fireproof gloves and clothing, and motor vehicle brake linings. Asbestos is a toxic material, and occupational exposure to asbestos dust can result in mesothelioma, asbestosis, and lung cancer.

The EPA began these proceedings in 1979, when it issued an Advanced Notice of Proposed Rulemaking announcing its intent to explore the use of TSCA "to reduce the risk to human health posed by exposure to asbestos." . . .

An EPA-appointed panel reviewed over one hundred studies of asbestos and conducted several public meetings. Based upon its studies and the public comments, the EPA concluded that asbestos is a potential carcinogen at all levels of exposure, regardless of the type of asbestos or the size of the fiber. The EPA concluded in 1986 that exposure to asbestos "poses an unreasonable risk to human health" and thus proposed at least four regulatory options for prohibiting or restricting the use of asbestos, including a mixed ban and phase-out of asbestos over ten years; a two-stage ban of asbestos, depending upon product usage; a three-stage ban on all asbestos products leading to a total ban in ten years; and labeling of all products containing asbestos.

Over the next two years, the EPA updated its data, received further comments, and allowed cross-examination on the updated documents. In 1989, the EPA issued a final rule prohibiting the manufacture, importation, processing, and distribution in commerce of most asbestos-containing products. Finding that asbestos constituted an unreasonable risk to health and the environment, the EPA promulgated a staged ban of most commercial uses of asbestos. The EPA estimates that this rule will save either 202 or 148 lives, depending upon whether the benefits are discounted, at a cost of approximately \$450–800 million, depending upon the price of substitutes.

\* \* \*

IV. The Language of TSCA.

\* \* \*

B. The EPA's Burden Under TSCA

\* \* \*

We conclude that the EPA has presented insufficient evidence to justify its asbestos ban. We base this conclusion upon two grounds: the failure of the EPA to consider all necessary evidence and its failure to give adequate weight to statutory language requiring it to promulgate the least burdensome, reasonable regulation required to protect the environment adequately. Because the EPA failed to address these concerns, and because the EPA is required to articulate a "reasoned basis" for its rules, we are compelled to return the regulation to the agency for reconsideration.

#### 1. Least Burdensome and Reasonable.

TSCA requires that the EPA use the least burdensome regulation to achieve its goal of minimum reasonable risk. This statutory requirement can create problems in evaluating just what is a "reasonable risk." Congress's rejection of a no-risk policy, however, also means that in certain cases, the least burdensome yet still adequate solution may entail somewhat more risk than would other, known regulations that are far more burdensome on the industry and the economy. The very language of TSCA requires that the EPA, once it has determined what an acceptable level of non-zero risk is, choose the least burdensome method of reaching that level.

In this case, the EPA banned, for all practical purposes, all present and future uses of asbestos — a position the petitioners characterize as the "death penalty alternative," as this is the most burdensome of all possible alternatives listed as open to the EPA under TSCA. . . .

By choosing the harshest remedy given to it under TSCA, the EPA assigned to itself the toughest burden in satisfying TSCA's requirement that its alternative be the least burdensome of all those offered to it. Since, both by definition and by the terms of TSCA, the complete ban of manufacturing is the most burdensome alternative — for even stringent regulation at least allows a manufacturer the chance to invest and meet the new, higher standard — the EPA's regulation cannot stand if there is any other regulation that would achieve an acceptable level of risk as mandated by TSCA.

\* \* \*

Much of the EPA's analysis is correct, and the EPA's basic decision to use TSCA as a comprehensive statute designed to fight a multi-industry problem was a proper one that we uphold today on review. What concerns us, however, is the manner in which the EPA conducted some of its analysis. TSCA requires the EPA to consider, along with the effects of toxic substances on human health and the environment, "the benefits of such substance[s] or mixture[s] for various uses and the availability of substitutes for such uses," as well as "the reasonably ascertainable economic consequences of the rule, after consideration for the effect on the national economy, small business, technological innovation, the environment, and public health."

The EPA presented two comparisons in the record: a world with no further regulation under TSCA, and a world in which no manufacture of asbestos takes place. The EPA rejected calculating how many lives a less burdensome regulation would save, and at what cost. Furthermore the EPA, when calculating the benefits of its ban, explicitly refused to compare it to an improved workplace in which currently available control technology is utilized. This decision artificially inflated the purported benefits of the rule by using a baseline comparison substantially lower than what currently available technology could yield.

Under TSCA, the EPA was required to evaluate, rather than ignore, less burdensome regulatory alternatives. TSCA imposes a least-to-most-burdensome hierarchy. In order to impose a regulation at the top of the hierarchy — a total ban of asbestos — the EPA must show not only

that its proposed action reduces the risk of the product to an adequate level, but also that the actions Congress identified as less burdensome also would not do the job. The failure of the EPA to do this constitutes a failure to meet its burden of showing that its actions not only reduce the risk but do so in the Congressionally-mandated least burdensome fashion.

Thus it was not enough for the EPA to show, as it did in this case, that banning some asbestos products might reduce the harm that could occur from the use of these products. If that were the standard, it would be no standard at all, for few indeed are the products that are so safe that a complete ban of them would not make the world still safer.

This comparison of two static worlds is insufficient to satisfy the dictates of TSCA. While the EPA may have shown that a world with a complete ban of asbestos might be preferable to one in which there is only the current amount of regulation, the EPA has failed to show that there is not some intermediate state of regulation that would be superior to both the currently-regulated and the completely-banned world. Without showing that asbestos regulation would be ineffective, the EPA cannot discharge its TSCA burden of showing that its regulation is the least burdensome available to it.

Upon an initial showing of product danger, the proper course for the EPA to follow is to consider each regulatory option, beginning with the least burdensome, and the costs and benefits of regulation under each option. The EPA cannot simply skip several rungs, as it did in this case, for in doing so, it may skip a less-burdensome alternative mandated by TSCA. Here, although the EPA mentions the problems posed by intermediate levels of regulation, it takes no steps to calculate the costs and benefits of these intermediate levels. Without doing this it is impossible, both for the EPA and for this court on review, to know that none of these alternatives was less burdensome than the ban in fact chosen by the agency.

The EPA's offhand rejection of these intermediate regulatory steps is "not the stuff of which substantial evidence is made." While it is true that the EPA considered five different ban options, these differed solely with respect to their effective dates. The EPA did not calculate the risk levels for intermediate levels of regulation, as it believed that there was no asbestos exposure level for which the risk of injury or death was zero. Reducing risk to zero, however, was not the task that Congress set for the EPA in enacting TSCA. The EPA thus has failed "cogently [to] explain why it has exercised its discretion in a given manner," by failing to explore in more than a cursory way the less burdensome alternatives to a total ban.

## 2. The EPA's Calculations.

Furthermore, we are concerned about some of the methodology employed by the EPA in making various of the calculations that it did perform. In order to aid the EPA's reconsideration of this and other cases, we present our concerns here.

First, we note that there was some dispute in the record regarding the appropriateness of discounting the perceived benefits of the EPA's rule. In choosing between the calculated costs and benefits, the EPA presented variations in which it discounted only the costs, and counter-variations in which it discounted both the costs and the benefits, measured in both monetary and human injury terms. As between these two variations, we choose to evaluate the EPA's work using its discounted benefits calculations.

Although various commentators dispute whether it ever is appropriate to discount benefits when they are measured in human lives, we note that it would skew the results to discount only costs without according similar treatment to the benefits side of the equation. Adopting the position of the commentators who advocate not discounting benefits would force the EPA similarly not to calculate costs in present discounted real terms, making comparisons

difficult. Furthermore, in evaluating situations in which different options incur costs at varying time intervals, the EPA would not be able to take into account that soon-to-be-incurred costs are more harmful than postponable costs. Because the EPA must discount costs to perform its evaluations properly, the EPA also should discount benefits to preserve an apples-to-apples comparison, even if this entails discounting benefits of a non-monetary nature.

When the EPA does discount costs or benefits, however, it cannot choose an unreasonable time upon which to base its discount calculation. Instead of using the time of injury as the appropriate time from which to discount, as one might expect, the EPA instead used the time of exposure.

The difficulties inherent in the EPA's approach can be illustrated by an example. Suppose two workers will be exposed to asbestos in 1995, with worker X subjected to a tiny amount of asbestos that will have no adverse health effects, and worker Y exposed to massive amounts of asbestos that quickly will lead to an asbestos-related disease. Under the EPA's approach, which takes into account only the time of exposure rather than the time at which any injury manifests itself, both examples would be treated the same. The EPA's approach implicitly assumes that the day on which the risk of injury occurs is the same day the injury actually occurs. Such an approach might be proper when the exposure and injury are one and the same, such as when a person is exposed to an immediately fatal poison, but is inappropriate for discounting toxins in which exposure often is followed by a substantial lag time before manifestation of injuries.

\* \* \*

Under the EPA's calculations, a twenty-year-old worker entering employment today still would be at risk from workplace dangers for more than thirty years after the EPA's analysis period had ended. The true benefits of regulating asbestos under such calculations remain unknown. The EPA cannot choose to leave these benefits high and then use the high unknown benefits as a major factor justifying EPA action.

\* \* \*

### 3. Reasonable Basis.

\* \* \*

Most problematical to us is the EPA's ban of products for which no substitutes presently are available. In these cases, the EPA bears a tough burden indeed to show that under TSCA a ban is the least burdensome alternative, as TSCA explicitly instructs the EPA to consider "the benefits of such substance or mixture for various uses and the availability of substitutes for such uses." These words are particularly appropriate where the EPA actually has decided to ban a product, rather than simply restrict its use, for it is in these cases that the lack of an adequate substitute is most troubling under TSCA.

As the EPA itself states, "[w]hen no information is available for a product indicating that cost-effective substitutes exist, the estimated cost of a product ban is very high." Because of this, the EPA did not ban certain uses of asbestos, such as its use in rocket engines and battery separators. The EPA, however, in several other instances, ignores its own arguments and attempts to justify its ban by stating that the ban itself will cause the development of low-cost, adequate substitute products.

As a general matter, we agree with the EPA that a product ban can lead to great innovation, and it is true that an agency under TSCA, as under other regulatory statutes, "is empowered to issue safety standards which require improvements in existing technology or which require the development of new technology." As even the EPA acknowledges, however, when no adequate substitutes currently exist, the EPA cannot fail to consider this lack when

formulating its own guidelines. Under TSCA, therefore, the EPA must present a stronger case to justify the ban, as opposed to regulation, of products with no substitutes.

\* \* \*

This presents two problems. First, TSCA instructs the EPA to consider the relative merits of its ban, as compared to the economic effects of its actions. The EPA cannot make this calculation if it fails to consider the effects that alternate substitutes will pose after a ban.

Second, the EPA cannot say with any assurance that its regulation will increase workplace safety when it refuses to evaluate the harm that will result from the increased use of substitute products. While the EPA may be correct in its conclusion that the alternate materials pose less risk than asbestos, we cannot say with any more assurance than that flowing from an educated guess that this conclusion is true.

Considering that many of the substitutes that the EPA itself concedes will be used in the place of asbestos have known carcinogenic effects, the EPA not only cannot assure this court that it has taken the least burdensome alternative, but cannot even prove that its regulations will increase workplace safety. Eager to douse the dangers of asbestos, the agency inadvertently actually may increase the risk of injury Americans face. The EPA's explicit failure to consider the toxicity of likely substitutes thus deprives its order of a reasonable basis.

\* \* \*

In short, a death is a death, whether occasioned by asbestos or by a toxic substitute product, and the EPA's decision not to evaluate the toxicity of known carcinogenic substitutes is not a reasonable action under TSCA. Once an interested party brings forth credible evidence suggesting the toxicity of the probable or only alternatives to a substance, the EPA must consider the comparative toxic costs of each. Its failure to do so in this case thus deprived its regulation of a reasonable basis, at least in regard to those products as to which petitioners introduced credible evidence of the dangers of the likely substitutes.

#### 4. Unreasonable Risk of Injury.

The final requirement the EPA must satisfy before engaging in any TSCA rulemaking is that it only take steps designed to prevent "unreasonable" risks. In evaluating what is "unreasonable," the EPA is required to consider the costs of any proposed actions and to "carry out this chapter in a reasonable and prudent manner [after considering] the environmental, economic, and social impact of any action."

As the District of Columbia Circuit stated when evaluating similar language governing the Federal Hazardous Substances Act, "[t]he requirement that the risk be 'unreasonable' necessarily involves a balancing test like that familiar in tort law: The regulation may issue if the severity of the injury that may result from the product, factored by the likelihood of the injury, offsets the harm the regulation itself imposes upon manufacturers and consumers." *Forester v. CPSC*, 559 F.2d 774, 789 (D.C.Cir.1977). . . .

That the EPA must balance the costs of its regulations against their benefits further is reinforced by the requirement that it seek the least burdensome regulation. While Congress did not dictate that the EPA engage in an exhaustive, full-scale cost-benefit analysis, it did require the EPA to consider both sides of the regulatory equation, and it rejected the notion that the EPA should pursue the reduction of workplace risk at any cost. . . .

Even taking all of the EPA's figures as true, and evaluating them in the light most favorable to the agency's decision (non-discounted benefits, discounted costs, analogous exposure estimates included), the agency's analysis results in figures as high as \$74 million per life saved. For example, the EPA states that its ban of asbestos pipe will save three lives over the

next thirteen years, at a cost of \$128–227 million (\$43–76 million per life saved), depending upon the price of substitutes; that its ban of asbestos shingles will cost \$23–34 million to save 0.32 statistical lives (\$72–106 million per life saved); that its ban of asbestos coatings will cost \$46–181 million to save 3.33 lives (\$14–54 million per life saved); and that its ban of asbestos paper products will save 0.60 lives at a cost of \$4–5 million (\$7–8 million per life saved). Were the analogous exposure estimates not included, the cancer risks from substitutes such as ductile iron pipe factored in, and the benefits of the ban appropriately discounted from the time of the manifestation of an injury rather than the time of exposure, the costs would shift even more sharply against the EPA's position.

While we do not sit as a regulatory agency that must make the difficult decision as to what an appropriate expenditure is to prevent someone from incurring the risk of an asbestos-related death, we do note that the EPA, in its zeal to ban any and all asbestos products, basically ignored the cost side of the TSCA equation. The EPA would have this court believe that Congress, when it enacted its requirement that the EPA consider the economic impacts of its regulations, thought that spending \$200–300 million to save approximately seven lives (approximately \$30–40 million per life) over thirteen years is reasonable.

\* \* \*

The EPA's willingness to argue that spending \$23.7 million to save less than one-third of a life reveals that its economic review of its regulations, as required by TSCA, was meaningless. As the petitioners' brief and our review of EPA caselaw reveals, such high costs are rarely, if ever, used to support a safety regulation. If we were to allow such cavalier treatment of the EPA's duty to consider the economic effects of its decisions, we would have to excise entire sections and phrases from the language of TSCA. Because we are judges, not surgeons, we decline to do so.<sup>8</sup>

\* \* \*

## VI. Conclusion.

In summary, of most concern to us is that the EPA has failed to implement the dictates of TSCA and the prior decisions of this and other courts that, before it imposes a ban on a product, it first evaluate and then reject the less burdensome alternatives laid out for it by Congress. While the EPA spent much time and care crafting its asbestos regulation, its explicit failure to consider the alternatives required of it by Congress deprived its final rule of the reasonable basis it needed to survive judicial scrutiny.

\* \* \*

Finally, the EPA failed to provide a reasonable basis for the purported benefits of its proposed rule by refusing to evaluate the toxicity of likely substitute products that will be used to replace asbestos goods. While the EPA does not have the duty under TSCA of affirmatively seeking out and testing all possible substitutes, when an interested party comes forward with credible evidence that the planned substitutes present a significant, or even greater, toxic risk

---

<sup>8</sup> As the petitioners point out, the EPA regularly rejects, as unjustified, regulations that would save more lives at less cost. For example, over the next 13 years, we can expect more than a dozen deaths from ingested toothpicks — a death toll more than twice what the EPA predicts will flow from the quarter-billion-dollar bans of asbestos pipe, shingles, and roof coatings. See L. Budnick, *Toothpick-Related Injuries in the United States, 1979 Through 1982*, 252 J. Am. Med. Ass'n, Aug. 10, 1984, at 796 (study showing that toothpick-related deaths average approximately one per year).

than the substance in question, the agency must make a formal finding on the record that its proposed action still is both reasonable and warranted under TSCA.

We regret that this matter must continue to take up the valuable time of the agency, parties and, undoubtedly, future courts. The requirements of TSCA, however, are plain, and the EPA cannot deviate from them to reach its desired result. We therefore GRANT the petition for review, VACATE the EPA's proposed regulation, and REMAND to the EPA for further proceedings in light of this opinion.

### Notes and Questions

**1. All Risk Is Relative:** Much of the court's opinion discusses the relative risk of asbestos alternatives. In many instances, there was evidence that suggested that the risk of substitute products outweighed the risk of asbestos. By ignoring the relative risks of different products, the EPA's rulemaking might have resulted in a net increase in lost lives. Why does the EPA choose to ignore such tradeoffs? Why isn't the EPA more concerned with regulating these substitute products?

The use of this kind of analysis was initially embraced by the EPA in a 1987 study entitled *Unfinished Business: A Comparative Assessment of Environmental Problems*. The study relied on an emerging technique known as **comparative risk analysis** (CRA). CRA is an economic tool used to quantitatively assess, compare, and rank different environmental and health risks. The study looked at over thirty health concerns and ranked them in terms of various risks including cancer, non-cancer, ecological, and welfare risks. A follow-up study by an EPA subcommittee published in 1990 entitled *Reducing Risk* found fault with many of the methods used in the original study but concluded that:

The Subcommittee developed a strong consensus that the relative risk assessment process is a good mechanism to formulate public policy from a scientific base of data and mechanistic processes. We recommend that the Agency institutionalize this approach on a regular basis, and provide the trained personnel and scientific data-bases needed to establish a scientific credibility for the process.

The use of this method remains controversial, and the academic community has split over the correct methods and utility of this approach. Do you think this is a good way to allocate scarce social resources to address health and safety risks? Could this method be usefully applied to other areas of public policy? Should other government agencies adopt the use of comparative risk analysis? Should private individuals?

**2. Cost-Benefit Analysis and Discounting:** In some cases, the benefits that may result from regulatory efforts are not realized until many years into the future. For example, environmental regulations often contemplate current expenditures for delayed benefits. In order to determine the relevant costs and benefits of such regulations, regulators must discount future benefits and costs to present values. The basic economic approach is that immediate benefits are more valuable than benefits in the future because of the likelihood of changing circumstances and the uncertainty of success. On numerous occasions, the EPA has argued that future benefits should not be discounted. But by not discounting, the benefits of EPA regulations appear to be much larger than they otherwise would be.

**3. Establishing an Unreasonable Risk:** The Toxic Substances Control Act (TSCA) contemplates that "unreasonable" risks will be regulated. In interpreting this statutory language, courts often refer to the reasonableness standard used in tort law. That is, if the severity of injury multiplied by the probability of injury is greater than the costs of regulation, then an



unreasonable risk is presumed to exist. Even under an interpretation of the evidence most favorable to the EPA, the court has doubts as to the need for any regulation. Do you agree with the court's conclusion in this regard? The court often gave dollar values for the cost of lives saved under the EPA regulation. What value was the court placing on the expected benefit of a life saved, and how do you think the court determined the value of a life to begin with? After considering the numerous methodological and procedural errors made by the EPA is it clear that an unreasonable risk exists? Was the EPA simply sloppy in quantifying the riskiness of asbestos? What other incentives might the EPA face which would cause it to make the type of errors discussed by the court?

**4. *The Least Burdensome Regulation:*** Once an unreasonable risk has been identified, the TSCA requires that the least burdensome approach to regulation be taken. The EPA provides information regarding the benefits and costs of essentially a complete ban on asbestos. The court's primary concern in this regard is that it does not have information as to whether a complete ban or some other level of regulation provides the largest net benefit. In other words, the court interprets the TSCA as requiring the most bang for the regulatory buck. Can you explain this standard in terms of marginal analysis? Is this an appropriate guide by which risk regulation should be approached?

Note that the court vacates the decision and remands the case back to the EPA for further proceedings. Seeing as the court was clearly skeptical of the EPA's analysis and found many of its conclusions lacking in support, is this the best remedy? What guarantees are there that the EPA will do a more thorough job the second time around? Is the only safeguard that the EPA can again be sued if the future form of this regulation continues to be insufficient?

**5. *Irrationality and Uncertainty:*** Empirical studies suggest that many individuals have difficulty conceptualizing probabilities. As a result, individuals tend to overestimate the risks of low probability events and underestimate the risks of high probability events. This result has a variety of implications for risk regulation. First, it suggests that even if market participants were supplied with full information, the optimal level of risk might not prevail. However, in the case of small probabilities such results would suggest the presence of overregulation. Second, to the extent that policy makers are responsive to public perceptions of risk, government regulation will not be optimal. Furthermore, studies suggest that individuals tend to overestimate the risks of events which are frequently reported on the news. Thus, even though the chance of being killed in a car wreck is substantially higher than death from AIDS, public perception often places the relative likelihoods of these two risks much closer together.

Consider an example. Steven Levitt, an economist at the University of Chicago, wrote an article noting that many parents may be hesitant to allow their children to play at the house of a friend whose parents own guns, but that they are rarely concerned with the presence of a pool. This assessment ignores the relative risks of those two factors. Levitt writes that:

What's more dangerous: a swimming pool or a gun? When it comes to children, there is no comparison: a swimming pool is 100 times more deadly.

In 1997 alone (the last year for which data are available), 742 children under the age of 10 drowned in the United States last year alone. Approximately 550 of those drownings — about 75 percent of the total — occurred in residential swimming pools. According to the most recent statistics, there are about six million residential pools, meaning that one young child drowns annually for every 11,000 pools.

About 175 children under the age of 10 died in 1998 as a result of guns. About two-thirds of those deaths were homicides. There are an estimated 200 million guns in

the United States. Doing the math, there is roughly one child killed by guns for every one million guns.

Thus, on average, if you both own a gun and have a swimming pool in the backyard, the swimming pool is about 100 times more likely to kill a child than the gun is.

Don't get me wrong. My goal is not to promote guns, but rather, to focus parents on an even greater threat to their children. People are well aware of the danger of guns and, by and large, gun owners take the appropriate steps to keep guns away from children. Public attitudes towards pools, however, are much more cavalier because people simply do not know the facts.

Steven D. Levitt, *Pools More Dangerous Than Guns*, Chicago Sun-Times, July 28, 2001.

### **3. Risk versus Risk**

Well-intentioned risk regulations often have the unintended consequence of increasing other types of risk. For example, empirical studies indicate that regulations requiring safety belts in cars changed drivers' attitudes towards speeding. In short, when safely buckled into their cars, many drivers felt that it was less risky to speed. The unintended consequence was to increase the risk to pedestrians and motorcyclists of being hit by a car. Thus, the safety belt regulation, while potentially saving the lives of those driving cars, inadvertently increased the risk to other groups of people being killed.

Perhaps the most basic error made by risk regulators in this regard is to forget that risk outcomes are a function of both the characteristics of the product or service in question and individual behavior. As a consequence, regulators often ignore the predictable rational economic responses of resourceful individuals. Recall the REMM model from Chapter I. Suppose that government regulators believed that requiring child resistant caps on cleaning products would reduce the risk of child poisoning from 1 in 100,000 children to 1 in 200,000 children per year. Before adopting this regulation, however, the agency should engage in **risk tradeoff analysis** (RTA). RTA forces regulators to go through the process of thinking about how reducing one form of risk impacts human behavior and consequently the presence of other types of risk — risk versus risk. For example, the child resistant cap regulation proposed above may have an impact on parents' perceptions of the marginal benefits of precautionary measures that they took to avoid child poisonings prior to the adoption of the regulation. For instance, because of the child safety caps, parents might be less diligent about placing cleaning products on high shelves, and some children might be poisoned because they have easier access to the cleaning products. Thus, the RTA suggests that some portion of the benefits derived from child safety caps might be offset by a reduction in individual care. The difficulty from the regulators' perspective is to quantify these risk tradeoffs in order to determine the net risk impact.

### **Competitive Enterprise Institute v. NHTSA**

United States Court of Appeals for the District of Columbia Circuit

956 F.2d 321 (1992)

and accompanying Notes and Questions

Read *supra*, Chapter II

### **4. Risk Regulation Priorities**

Risk reduction is costly. Every decision to reduce a particular risk therefore entails

opportunity costs — the opportunity to reduce some other form of risk. Thus, decisions need to be made regarding which risks will be attacked through regulation. The analysis in preceding subsections suggests several guidelines that should be kept in mind when setting these priorities. First, risk reduction at a point where marginal costs are greater than marginal benefits not only increases risk by reducing the size of the economic pie, but also takes resources away from cost-effective risk reduction efforts. Second, risk-tradeoff analysis suggests that regulations with the largest net reduction in risk per dollar spent should be pursued first. In other words, regulations that save lives at lower costs should take priority over those that save lives at higher costs. An empirical study by Dr. Tammy Tengs suggests that the current regulatory approach ignores these factors:

Results indicate that we incur opportunity costs of approximately \$31.1 billion, 60,200 premature deaths, or 636,000 years of life lost every year in order to maintain our present pattern of investment in these 185 life-saving interventions. At our current level of resource consumption, we could double the survival benefits of our expenditures. Alternatively, we could retain our present level of risk reduction and, in addition, save billions of dollars.

Tammy O. Tengs, *Optimizing Societal Investments in the Prevention of Premature Death*, doctoral dissertation, Harvard School of Public Health 2 (June 1994).

In the context of determining the appropriate priorities for risk regulation, it is important to recognize that regulation does not operate within a vacuum. Specifically, both the market and the legal system have an impact on prevailing safety levels. Thus, the prioritization question should not be limited to the cost effectiveness of different regulations, but rather should include the relative cost effectiveness of all sources of risk reduction. For example, empirical evidence on compensating wage differentials suggests that workers are generally capable of identifying on the job safety hazards, such as working with dangerous machinery. On the other hand, workers have a more difficult time identifying health hazards, such as the carcinogenic risk due to on the job toxins. Despite these facts, OSHA has focused its emphasis primarily on regulating safety hazards.

### **5. Political Economy of Health and Safety Regulation**

Cost-benefit analysis suggests that many risk regulations are too strict, yet government regulators continue to promulgate such regulations. Rather than assuming that the regulators are either ignorant, evil, or both, it might be helpful to consider the political incentives of those responsible for the oversight of the regulatory process. In general, risk regulators' incentives push them to adopt regulations that ignore the efficient determination of optimal risk levels. Three specific sources of these perverse incentives should be noted.

First, the literature on the economics of bureaucracy suggests that bureaucrats are motivated to maximize the size of their budget. In this regard, risk regulators have an incentive to increase the size and budget of their agency by using command and control methods of regulation rather than encourage the most efficient means for risk reduction. Often times, regulators attempt to reduce risk by imposing a particular set of technological requirements on an industry. This provides the regulatory agency with a large degree of control over how risk reduction is achieved. The agency needs a larger budget in order to control activity.

Second, both politicians and regulators tend to be risk averse because negative news impacts their personal interest more than positive news. Consider the case of a new drug that is expected to save 1,000 lives within its first year. Unfortunately, 100 individuals are expected to

die because of adverse side effects. Politicians and bureaucrats reasonably expect that the 100 killed will be bigger news than the 1,000 saved. When budgets and re-election are a function of political popularity, very few individuals will want to be known as the regulator or politician who allowed a new drug to go on the market that killed 100 people. Understandably, bureaucrats and politicians have an incentive to seek higher than optimal levels of risk reduction in order to maintain their jobs or their political status. Furthermore, to the extent that politicians respond to voter preferences and voters overestimate the risk of low probability events and underestimate the risk of higher probability events, an inefficient level of risk regulation is likely to result.

Finally, risk regulators and politicians often ignore the costs that they impose on others. The cost-benefit analysis of risk regulations clearly indicates that many areas are being regulated too intensively. More intelligent priority setting — that is, reallocation of risk reducing resources from one area to another — seems to be called for. However, it's important to recognize that the cost-benefit analysis compares the private cost of one industry with the benefits to society. If a particular industry is granted regulatory relief, the money saved by that industry is not available to regulators to use more wisely in regulating other industries — the savings belong to the first industry. Thus, regulators have little to gain by reducing regulation to the optimal level.

**Table VI-1. Cost per Life Saved for Arsenic Regulation**

<b>Stringency</b>	<b>Standard Level (mg/m<sup>3</sup>)</b>	<b>Average Cost Per Life Saved</b>	<b>Marginal Cost Per Life Saved</b>
Loose	0.10	\$1.25 million	\$1.25 million
Medium	0.05	\$2.95 million	\$11.5 million
Tight	0.004	\$5.63 million	\$68.1 million

Source: W. Kip Viscusi, *Risk by Choice: Regulating Health and Safety in the Workplace* 124 (1983)

Table VI-1. Cost per Life Saved for Arsenic Regulation

<b>Stringency</b>	<b>Standard Level (mg/m<sup>3</sup>)</b>	<b>Average Cost Per Life Saved</b>	<b>Marginal Cost Per Life Saved</b>
Loose	0.10	\$1.25 million	\$1.25 million
Medium	0.05	\$2.95 million	\$11.5 million
Tight	0.004	\$5.63 million	\$68.1 million

Source: W. Kip Viscusi, *Risk by Choice: Regulating Health and Safety in the Workplace* 124 (1983)

Figure VI-1. Utility Curve Exhibiting Diminishing Marginal Utility of Money Value

Figure VI-2. Decision Making under Uncertainty for Risk Averse Individuals

Figure VI-3. Risk Neutral

Figure VI-4. Risk Seeker

Figure VI-5. Market Determination of Workplace Safety

Figure VI-6. The Hand Formula