From: Advances in Applied Micro-Economics, edited by V. Kerry Smith, Volume 4, 1986, "Risk, Uncertainty, and the Value of Benefits and Costs." JAI Press, Inc., Greenwich, Conn., and London.

# THE DANGERS OF CAUTION:
## CONSERVATISM IN ASSESSMENT AND THE MISMANAGEMENT OF RISK

Albert L. Nichols and Richard J. Zeckhauser


## ABSTRACT

Regulatory agencies faced with major scientific uncertainties in areas ranging from nuclear power to toxic wastes frequently employ conservative assumptions that impart an upward bias to their risk assessments. This approach confuses risk assessment with risk management, distorting science to influence policy. Although motivated by a desire to emphasize safety over other considerations, such as cost, this bias leads to an inefficient allocation of risk-reduction resources because the degree of conservatism varies widely across risk assessments. If this inefficiency is sufficiently severe, conservatism can lead to less rather than more safety.

This paper examines the problems created by conservative assessments,


Advances in Applied Micro-Economics, Volume 4, pages 55–82.

ISBN: 0-89232-577-1

focusing on carcinogens. It examines the different stages at which conservative assumptions can be introduced, and shows how the cumulative effect of such assumptions can be far larger than is commonly realized. It then shows how asymmetries in conservatism (e.g., between the treatment of carcinogenic and noncarcinogenic risks) produce misallocations of control efforts. An analysis of the debate about the Environmental Protection Agency's (EPA's) recently tightened limits on lead in gasoline illustrates these potential problems. The paper argues that decision makers should rely on expected values of the estimated probabilities of losses (e.g., lives lost or cases of cancer), not on conservative estimates. The desire to give great emphasis to safety should be reflected in the decision rules applied by risk managers, not in distorted probability estimates. The paper identifies several steps to improve the realism of assessments.

## I. INTRODUCTION

Conservatism in risk assessment is the standard approach in a wide range of regulatory decisions, in fields ranging from control of toxic wastes to approval of new drugs. Motivated by a desire to emphasize safety over other considerations, the conservative approach employs assumptions that tend to overstate risk. This bias in assessment leads to an inefficient allocation of risk-reduction resources. The outcome even may be less safety than would result from decisions based on more realistic assumptions.

The fundamental problem is that conservatism confuses risk assessment with risk management. In its 1983 report, the National Academy of Sciences' Committee on Institutional Means for Assessment of Risks to Public Health recommended that regulatory agencies distinguish between "risk assessment," the primarily scientific enterprise of estimating the hazard associated with a substance or activity, and "risk management," the design and implementation of policies to control risks. Risk management relies primarily on political, ethical, and economic judgments informed, but not determined, by science. It involves such decisions as what levels of risk are acceptable and what trade-offs are appropriate between risk and other factors, such as control costs (National Academy of Sciences, 1983).

During his second tenure as administrator of the Environmental Protection Agency (1983–1985), William Ruckelshaus embraced the assessment–management distinction as an important organizing principle. He emphasized the need to improve the scientific quality and consistency of EPA's risk assessments. On the management side, he argued for procedural and statutory changes to help decision makers give more comprehensive and consistent consideration to the trade-offs in environmental decisions.[1]

Separating assessment and management offers important advantages. It shields the scientific process from political manipulation; decision makers cannot alter the risk assessment to fit the policy choice they wish to make.



Conversely, decision makers do not have to accept implicit policy choices masquerading as scientific fact. The distinction can clarify public discussion and accountability by making the scientific and value judgments that underlie decisions more explicit.

In practice, the line between risk assessment and management is often blurred. Fundamental gaps in scientific knowledge and data limitations make risk assessment a highly uncertain endeavor, requiring many choices among competing models and assumptions. There is a strong temptation to have such choices reflect implicit policy judgments rather than science.

This blur is most apparent in current techniques for estimating the risks associated with carcinogens, which employ conservative assumptions that bias the estimates upward. The intent is to err on the side of safety by minimizing the chance that risks will be underestimated and thus undercontrolled. But this approach intrudes the risk assessment process into risk management. In deciding how conservative to make their estimates, risk assessors implicitly trade off risk against other factors. Unless they explicitly acknowledge these trade-offs and quantify them, their assessments will mislead others, including those charged with managing risks. For example, risk managers will be more likely to impose a costly regulation if they mistakenly believe that it can prevent ten cases of cancer than if they correctly realize that it will eliminate only one case.

Advocates of the conservative approach are likely to view this tendency toward greater control as a virtue; conservatism is intended to give extra weight to protecting public health and, under conditions of massive uncertainty, to err on the side of safety. In fact, however, conservative risk assessment is a deeply flawed approach to protecting public health. It violates the distinction between risk assessment and risk management, concealing value judgments and policy choices under a cloak of science. It creates capricious differences in the degrees of safety provided across different substances and policy areas, because degrees of conservatism vary widely. Finally, because regulators must make complex trade-offs among different risks (not only between risk and cost), conservatism can lead to less rather than more safety, by misdirecting public concern and scarce agency and societal resources.

We first show how traditional procedures for assessing carcinogenic risks inject conservatism at multiple stages. We then discuss how these procedures can distort risk management decisions and argue that appropriate risk assessments will generate expected values rather than upper-bound estimates.

## II. SOURCES OF CONSERVATISM IN RISK ASSESSMENT

Estimating the impact of a regulation on risk is a complex, multistage process. For expositional purposes, consider a highly simplified framework

with three stages—emissions, exposure, and risk—that correspond roughly to the division of labor among different disciplines.[2] Engineers may estimate how a regulation will affect emissions, while specialists in dispersion modeling translate emissions into exposure estimates, and biomedical experts and statisticians prepare the dose–response estimates that convert exposure to the outcome measure of ultimate concern, reductions in risk.

The overall reduction in risk due to a regulation is a multiplicative function of the individual elements: reduction in emissions, the amount of exposure per unit of emissions, and the level of risk per unit of exposure.[3] As a result, the individual uncertainties are compounded in the uncertainty attached to the overall estimate. Conservatism at each step enters multiplicatively in the final estimate.

## A. Emissions

The first step of the process is the most straightforward: estimating the levels of emissions that would prevail with and without a proposed regulation. The difference between these two quantities, both likely to be at least somewhat uncertain, represents the effect of the regulation. Although the following discussion refers to emissions—a term often reserved within EPA for air pollutants—it applies equally well to releases of pollutants into other media.

### *Baseline Emissions*

Baseline (preregulation) emissions depend primarily on emission rates (per unit of output) and the level of production. Emission estimates may be inflated by assuming that plants operate at full capacity. Often emission rates must be estimated from an engineering analysis of a model plant, though actual "uncontrolled" emission rates are likely to vary widely, depending on a host of source-specific factors. In addition, some sources already may have controls because of preexisting state or local regulations or economic self-interest (e.g., the recovery of valuable feedstocks). Regulatory agencies often base their priorities initially on old estimates of emission rates that are not representative of current technology and controls. Later analyses, available when regulations are proposed and promulgated, are likely to be more accurate, but by then some important choices may have been determined by the earlier, conservative assumptions. In setting priorities for regulating sources of benzene emissions, for example, EPA assigned high priority to maleic anhydride plants on the basis of emission estimates that later proved too high by an order of magnitude.[4]



### *Postregulation Emissions*

The effectiveness of a regulation may be quite uncertain, particularly when it is a design rather than a performance standard. How long, for example, will a particular liner prevent leaks from a hazardous-waste landfill? Similarly, the standards likely to be proposed soon for coke ovens are stated in such terms as the percentage of doors leaking; because the link between that measure and emissions is uncertain, the "minimum" and "maximum" emissions estimates vary by about an order of magnitude.[5]

Even when performance standards are contemplated, emission levels may be difficult to predict. To ensure that they will be in compliance, firms concerned about the excessive cost of having to retrofit if they fail to meet a standard may design controls to do better than required. Some firms may switch to a different production process, eliminating emissions entirely. Conversely, equipment failures, poor maintenance, or lax enforcement may lead to higher-than-expected emissions after the standard is imposed.[6]

## B. Exposure

The first step in estimating the link from emissions to exposure is typically to predict the dispersion patterns of the substance, generally by using computer models. For air pollutants, estimated concentrations based on the dispersion modeling are then combined with population data to estimate total population exposure and its distribution. For other pollutants, additional steps may be required, such as estimates of drinking water use (in the case of water pollutants) or of food consumption (for pesticide residues or contaminated fish). Each step is uncertain, requiring choices among alternative assumptions and estimates; thus each offers opportunities for (and temptations toward) conservatism.

### *Dispersion Modeling*

The accuracy of dispersion models varies widely across media and substances, but even the best models for the most tractable pollutants are uncertain (Miller, 1978). Often accuracy is further attenuated by the use of standard inputs to the models, rather than site-specific data. The models are less accurate when chemical reactions in the ambient environment are important.

Conservatism may enter dispersion modeling in the selection of parameter values. For air modeling, meteorological data may be used that lead to high ground-level concentrations (and thus higher exposure levels). Dispersion modeling for benzene emissions from maleic anhydride plants, for example, used data from Pittsburgh, where "meteorological conditions that maximize ground level concentrations . . . are common" (U.S. Environmental Protection Agency, 1980, pp. 4–11). A later modeling effort com-

missioned by the Chemical Manufacturers' Association, using essentially the same model but local meteorological data for each plant, predicted substantially lower concentrations (Galluzzo and Glassman, 1980).

For pollutants in the ground, conservative assumptions may be made about soil types and other key conditions (such as the presence of underground water flows), leading to faster and longer transport of the pollutants. Such conservatism is most likely when the analysis is restricted to a single model source.[7]

### *Population*

Estimates of ambient concentrations must be combined with data on human populations and behavior to calculate exposures. With air pollutants, for example, the models predict outdoor concentrations, but most people spend the vast majority of their time indoors. Recent studies have found little correlation between indoor and outdoor concentrations for many pollutants. Although indoor concentrations are higher because of indoor sources (which typically are not regulated by EPA), reductions in outdoor levels often have a less than one-for-one impact on indoor air (Spengler and Sexton, 1983).

For all types of pollutants, behavioral assumptions are important. Calculations of maximum individual exposure levels typically assume that some people are born and die at the point of maximum concentration, and never leave it, even to go to work or school. Estimates of the risk from groundwater contamination often fail to recognize that contamination may be discovered and alternative water supplies substituted. In general, failure to recognize that people may take steps to avoid exposure (for example, by buying bottled water if their well water is contaminated) leads to overestimates of exposure, and hence risk.[8]

## C. Dose–Response Relationship

Estimating the dose–response relationship that converts exposure to risk is by far the most uncertain and controversial step, with the greatest potential for the introduction of conservative assumptions. The primary sources of data are high-dose studies of laboratory animals and epidemiological studies of workers exposed to much higher concentrations than the general population would encounter. The major uncertainties in extrapolating from animals to humans and from high to low doses both provide fertile ground for conservatism.

### *Animal to Human*

Where possible, risk assessments rely on human epidemiological data. Often, however, only studies of laboratory animals are available. There



are obvious conceptual difficulties in extrapolating from small, genetically homogeneous animals (often bred to be particularly susceptible to cancer) to much larger, genetically diverse human beings, who live far longer. Although most scientists accept demonstrated carcinogenicity in animals as strong qualitative evidence that the substance will be carcinogenic in humans, methods for estimating human dose–response relationships from animal data remain highly controversial. Some scientists dispute the basic premise that data from small rodents provide useful starting points for quantitative predictions of human responses[9]; unfortunately, there is often no realistic alternative.

The sources of uncertainty are numerous, including general problems of quality control (particularly in older studies), differences in pathologists' interpretations of slides, differences in how long the animals are kept alive, and the fact that animal studies often involve different routes of exposure than those for human beings. Risk estimates for an inhaled airborne substance, for example, may be based on animal feeding studies.

The most concrete quantitative issue is the computation of equivalent doses. The two most widely used bases for translating animal doses to human equivalents are weight and surface area. This is no trivial issue. Because surface area increases much less, proportionally, than weight does as one moves from small to large animals, the surface–area conversion leads to much higher risk estimates in extrapolating from animals to humans. For mice, the difference between the two methods is roughly a factor of 13; for rats the difference is approximately a factor of 6.[10] Statistical studies of carcinogens that have been tested in more than one species fail to provide much basis for choosing among the two methods (Crouch and Wilson, 1979, 1981). EPA relies on the more conservative surface–area method.

Two other standard procedures also reflect a tendency to select the most conservative of several plausible alternatives. First, where multiple species (or strains of the same species) and sexes have been tested, data from the species-strain-sex group yielding the highest risk estimate are used (unless that study is unusually small or otherwise flawed). The primary rationale is that because human beings are genetically diverse, the most sensitive animal strain should be used to account for more sensitive segments of the human population. But the fact that some people are less sensitive than average, and others merely average, also should be considered in estimating population risks. In general, we want to know the average risk in the exposed population.[11]

Another source of conservatism is the counting of both benign and malignant tumors in animal studies. The argument is that benign tumors may progress to malignancy, and that a substance producing benign tumors in animals might produce malignant ones in human beings. Most benign tu-

mors, however, never become malignant, so this procedure almost certainly overestimates the incidence of cancer. To calculate the incidence of cancer in human beings in an analogous manner would mean autopsying everyone who died and counting all tumors, benign and malignant, whether or not they were the proximate cause of death; the resulting "cancer rate" would be much higher than the true one.

*Extrapolation to Low Doses*

Neither epidemiological data nor laboratory experiments with animals can detect low-level risks; the sample sizes needed to detect small increments in risk are simply too large (Schneiderman et al., 1975). Thus, risks to the general population ordinarily are assessed by studying individuals with unusually high exposures (often workers) or giving animals doses far beyond those likely to be encountered by people. Various mathematical models have been developed to perform the necessary extrapolations. Unfortunately, current theory does not provide unambiguous support for any one of them, nor can empirical tests determine which is most accurate.

Regulatory agencies most often use some variant of the "one-hit" model. At low doses, this model is very closely approximated by a straight line through the dose-risk origin; that is, it predicts that risk is proportional to dose. Thus, for example, if the relevant dose is 1000 times lower than the level at which the risk is measured, the estimated risk also is 1000 times lower. In estimating the dose–response function from epidemiological data, EPA's Carcinogen Assessment Group (CAG) usually uses the linear approximation of the one-hit model (Anderson, 1982).

With animal data, the CAG uses the multistage model, which is closely approximated by a polynomial at low doses (Anderson, 1982). (In those rare cases where epidemiological studies yield several dose groups, the CAG also uses this model for human data.) The number of nonlinear terms estimated is 1 less than the number of dose groups in the relevant study; for example, if the study has two dose groups in addition to the controls, coefficients are estimated for both dose and dose–squared. The CAG first computes the maximum likelihood estimates (MLEs) for each parameter, with the constraint that they all be nonnegative. In most cases, the linear term is nonzero, and it dominates the predicted risk at low doses. In some cases, however, the MLE of the linear term is zero. To force linearity at the low-dose level, in all cases (even when the MLE of the linear term is positive) the CAG uses the 95 percent upper confidence limit of the linear term, rather than the MLE, in making its risk estimates. When the MLE of the linear term is positive, this procedure increases the low-dose risk estimate by a factor of roughly 2 to 3. When the MLE of the linear term is zero, however, this procedure can increase the low-dose risk estimate



by several orders of magnitude.[12] Thus, although the CAG uses a model for animal data that allows for nonlinearity, its method of estimation ensures that the linear portion will dominate.[13]

The CAG's rationale for forcing linearity is twofold. First, the theory underlying the multistage model argues that the linear term should be positive. Second, low-dose predictions of risk based on maximum likelihood estimates of the multistage model can be extremely unstable with certain data sets. Two sets of parameter values, for example, one with a zero linear term and the other with a positive one, may yield virtually identical values of the likelihood function for the observed data, yet generate very different low-dose risk estimates. Moreover, in such cases, very small changes in the data (e.g., a change in the response of one animal in the experiment) can determine whether the linear term is positive, and thus generate low-dose risk estimates that vary by orders of magnitude. The 95 percent upper confidence approach does not suffer from this hypersensitivity.

Although reliance on a pure MLE approach probably would be undesirable, it should be possible to develop alternative procedures that preserve the desirable element of stability without unnecessarily injecting additional conservatism; as noted above, the current method increases the risk estimates by a factor of about 2 even when the MLE yields a positive linear term. EPA is now considering such alternatives.[14]

Virtually all of the other extrapolative models are convex at low doses; as the dose is reduced, risk falls more than proportionally. (Well-known nonlinear alternatives include the log-probit, logit, and multihit models.) With a two-hit model, for example, reducing the dose by a factor of 1000 reduces risk by a factor of roughly 1 million.[15] Thus, when estimated from the same high-dose data, these nonlinear models predict far smaller low-dose risks than the one-hit model or the linearized version of the multistage model.

Unfortunately, the various models often give similar fits to the data from laboratory or epidemiological studies, so empirical tests are inconclusive. When the extrapolation from data to the exposures of concern covers doses that differ by a factor of 100 or more (which typically is the case with environmental carcinogens), the nonlinear models usually predict risks that are hundreds or thousands of times lower than those predicted by the linear model.[16] Thus, with rare exceptions for extremely potent carcinogens or unusually high exposure levels, regulations to reduce exposure of the general population to environmental carcinogens do not seem justified unless one believes that the linear assumption has a reasonable probability of being correct (or at least provides a good approximation of the true dose–response function).[17]

Most scientists believe the one-hit model (and its variants) represents an upper-bound estimate of low-dose risks. Elizabeth Anderson (1982), then director of EPA's Office of Health and Environmental Assessment, noted that "the linear, nonthreshold curve is regarded by most scientists working in risk assessment in the United States, as usually placing a plausible upper-bound on risks, i.e., the risks are not likely to be higher." The CAG has been quite explicit that its estimates are meant to establish plausible upper bounds on risk, not to assess actual risk as precisely as possible. The group's reports generally acknowledge that the plausible lower bound is zero, but summary documents prepared for decision makers or the public often do not include such statements. Typically, numerical calculations (e.g., cost–effectiveness ratios) rely solely on the upper-bound estimate.

There is little agreement about the margin by which the linear assumption overpredicts risk. Some scientists seem to accept it as giving a plausible estimate of *actual* risks, though their support for the model may be based largely on a policy judgment that assessments should err on the high side. Other scientists believe that thresholds apply to carcinogens as well as to other substances, so that low-level exposures almost certainly are harmless.[18]

Recent work in biokinetics suggests that for many carcinogenic substances it is metabolites, rather than the substances themselves, that interact with DNA (deoxyribonucleic acid) to cause cancer. Even if the relationship between those metabolites and risk is linear, the relationship between exposure to the substance and risk is likely to be nonlinear, because the metabolic processes involved are nonlinear. For the most part, this suggests that risks at low doses will be substantially lower than those estimated by the linear model, although in a few cases the metabolic nonlinearity runs in the opposite direction, yielding higher risks at low doses than those predicted by the linear model (Hoel et al., 1983).[19] These studies also suggest that the shape of the dose–response function is likely to vary across substances depending on the specific metabolic processes involved, so any research for a single "right" functional form may be misguided.

### *An Example*

A decision-analytic case study of perchloroethylene (PCE) by Campbell et al. (1982) illustrates some of these issues and the ways in which conservative assumptions can affect the results. The authors consider two possibilities for each of three steps in estimating the unit risk factor for PCE: (1) the form of the dose–response model (linear vs. nonlinear); (2) the method of extrapolation from animals to human beings (surface area vs. weight); and (3) the species used for extrapolation (mouse vs. rat, where the mouse study showed a higher response rate). When eight ($2^3$) possible combinations of assumptions were evaluated, the resulting estimates varied by a factor of 35,000, ranging from the low estimate derived from a nonlinear, weight-based extrapolation from the rat study to the high estimate derived from a linear, surface-area-based extrapolation from the mouse study. Given current scientific knowledge, one cannot determine which set of assumptions gives the best estimate.[20] Standard risk-assessment procedures used by EPA and other agencies would use the most conservative assumption for each step, and thus would yield the highest estimate.



## III. CONSERVATISM AND MISALLOCATION

Conservative risk assessment procedures for carcinogens are the product of great scientific uncertainty and an implicit policy judgment that regulators should err on the side of safety. Few would disagree that risk assessment is beset by tremendous uncertainties or that EPA and other risk-management agencies should give great weight to safety in their decisions. But deliberate, exaggerated conservatism in risk assessment is an exceedingly inefficient way of erring on the side of safety and may even lead to less protection than would result from more realistic risk assessments.

Conservatism has many flaws as a way of providing a margin of safety. We believe decision makers who wish to err on the side of safety would do better to use unbiased risk estimates and then attach a high value to the risks reduced. Compared to the current approach, such a strategy would lead to less stringent regulation in some circumstances (those where current assessments are most conservative), but tighter regulation in other cases. Whether it would increase or decrease overall risk-reduction efforts is an open question; it is quite possible that expenditures and overall risk both would go down.

In this section we first show overall conservatism in carcinogen risk estimates may be far greater than commonly realized because conservative assumptions are employed at many points. Moreover, the degree of conservatism can vary widely across substances, because of inconsistencies in procedures or because the same procedure has very different implications for different substances and regulatory contexts. We then show how these differences in conservatism can lead to regulatory decisions and priorities that yield less safety than would be achieved with unbiased assessment procedures. Although our focus is on risk assessment, our arguments against conservatism apply as well to estimates of the dollar and other costs of controls, such as losses of jobs. The information provided decision makers about such costs should reflect realistic assumptions and the range of possible outcomes, not worst-case assumptions that overstate the likely economic impacts of regulation.

### A. Greater-than-Intended Conservatism

Assessing the risks posed by carcinogens is a highly uncertain enterprise. Assessors must choose among alternative assumptions and models on the basis of limited evidence. The general, though not universal, tendency is to make conservative (or worst-case) choices throughout the process. This tendency is likely to be reinforced by the fact that different groups are responsible for different stages. Each choice, when viewed in isolation, may appear reasonable and prudent. Multiplied together, however, even modest overestimates can easily compound to yield a substantial overestimate of the overall risk.

Suppose, for example, that the risk assessment involves five key steps, and that the final risk estimate is the product of those five steps. The steps might be emission level, exposure per unit of emissions, the method for converting from animal to human doses, the dose–response model, and the statistical method for estimating the model's parameters. Let the degree of conservatism applied at each step be relatively modest; experts estimate that each parameter value is the 90 percent upper confidence limit, and that it is twice the expected value. The cumulative effect is severe. Assuming that the individual steps were statistically independent, the final estimate would exceed its expected value by a factor of $2^5 = 32$ and would have far more than a 90 percent chance of being higher than the true value. Precisely how much higher depends on the density functions associated with each parameter estimate. If all the parameters were distributed log-normally with the same variances and were independent, the product of five 90th percentile estimates would have a 99.8 percent chance of being too high.[21]

In practice, decision makers typically cannot tell how conservative the final estimate is. Information on the conservative assumptions may not appear in the summary documents they see, and in any case the effects of the conservative assumptions rarely are quantified. As a result, the decision maker does not know whether the overall estimate in any particular case represents the 80th, 95th, or 99.9th percentile of the distribution of possible estimates. The estimate could be higher than the expected value by a factor of 2, 10, or 1000. The decision maker may not have even a rough qualitative feel for the degree of conservatism.[22]

### B. Variations in Conservatism

The degree of conservatism is likely to vary significantly from one risk assessment to another. Different agencies may follow inconsistent procedures for carcinogenic risk assessment. The Office of Science and Technology Policy (49 *Federal Register* 21594, 1984) has issued a set of cancer risk-assessment "principles" to promote greater consistency across agencies, but they leave much room for differences among agencies in the details of implementation. Consistency can be a problem even within a given agency; the risk assessment guidelines recently proposed by EPA (49 *Federal Register* 46294, 1984) represent an effort to ensure uniform procedures across the offices in that agency.



Uniform guidelines eliminate only one source of inconsistency due to conservatism. If all estimates that rely on animal tests, for example, use the results from the most sensitive species tested, the degree of conservatism will tend to rise with the number of tests conducted for a particular substance. Similarly, the practice of counting benign as well as malignant tumors in animal tests may distort relative-risk estimates if substances differ in the mix of tumors they cause in test animals. If the techniques used for animal tests are more conservative than those used when human epidemiological data are available (e.g., because of the factor used to convert animal doses to their human equivalents, the use of the most sensitive group tested, or the use of upper-confidence-limit parameter estimates), then the upward bias will be higher for substances for which the only data come from animal tests.

The assumption of linearity in the dose–response relationship also can impose different degrees of conservatism. Suppose that the available data are from an epidemiological study of farm workers exposed to high doses of a pesticide. Even if the true relationship is nonlinear, little bias will result from using the linear model with those data to predict the risk for other workers exposed to similar or slightly lower levels. (Indeed, any bias may be toward underestimating the amount of risk reduction achieved by controls, because nonlinear models show higher *marginal* risk at high doses.) But a substantial overestimate may result from using that same model and the same data to predict risks for members of the general population exposed to trace residues.[23]

### C. Misordered Priorities

Differences in the conservatism of assessments, even if all assessments are "too high," can lead to less rather than more safety. Suppose the true carcinogenic risk of substance A is twice that of substance B. Because of differing degrees of conservatism, however, the risk from A is overestimated only by a factor of 2, while the risk from B is overestimated by a factor of 10. Decision makers reviewing these estimates will conclude that B is more dangerous and deserves more stringent control than A, when in fact the reverse is true. The problem is particularly striking when the two substances are substitutes (e.g., alternative pesticides for the same use), and the decision maker must decide which one should remain in use. More generally, regulatory agencies have limited resources and can pursue only some of the actions they might like to undertake. EPA's Office of

Toxic Substances, for example, has identified thousands of potentially hazardous chemicals in use, but cannot afford to examine more than a handful at a time. The Office of Pesticide Programs faces similar problems in dealing with the backlog of reviews of existing pesticides.

### *Carcinogens vs. Noncarcinogens*

Not every pollutant is a carcinogen. The procedures for estimating noncarcinogenic risks typically are less conservative than those for carcinogens; that is, they lead to less extreme overestimates. Most often EPA does not attempt to quantify noncarcinogenic risks at all, focusing instead on ambient standards or on the technological feasibility of controls. When quantitative estimates are attempted, EPA usually does not extrapolate far beyond the range of exposures at which adverse effects have been found.[24]

As a result, priorities between carcinogens and noncarcinogens may get distorted, both in the context of individual regulations and within or across programs. Control of one pollutant may increase the level of another. Reducing the amount of lead in gasoline, for example, may increase the amount of benzene. Control of a pollutant in one medium may increase its presence in another; certain types of pollution treatment facilities, for example, eliminate toxic substances from water by volatilizing them to the air. If different degrees of conservatism are applied to different types of substances and media, decision makers may misperceive the true implications of the alternatives and take actions that increase overall risk.

Direct trade-offs between carcinogenic and noncarcinogenic risk are fairly unusual, but agency decision makers routinely must decide how to divide limited contract dollars and personnel between carcinogenic air pollutants and "conventional" (i.e., noncarcinogenic) ones. If they base such decisions on assessments of relative risks, differential conservatism can lead them astray, causing them to devote relatively too many resources to carcinogens and too few to noncarcinogens.

The potential distortion is seen most readily by comparing agencies charged with regulating very different types of risks. Some risks are common enough and easily enough measured that reliable statistics are available. We have very accurate counts of how many people die in motor vehicle accidents each year, for example. Predicting the effects of intervention — such as air bags — is more difficult, but still much less uncertain than predicting the health impact of lowering the emissions of a carcinogenic substance. As a result, conservatism is not a very important issue in risk assessments for such problems. Observers may argue about which estimate is more accurate, but the band of uncertainty is relatively narrow.[25] And because we have good data on the baseline risks, gross overestimates of risk reduction can be detected; the number of highway fatalities, for example, places a firm upper bound on estimates of the number of lives that might be saved by air bags.



Not so with carcinogenic risk assessment. The greater the uncertainty about a given effect, the more likely it is to be overestimated under conservative approaches.[26] As a result, decisions based on conservative assessments are likely to overcontrol highly uncertain risks, such as environmental carcinogens, relative to those that can be estimated with more precision, such as highway accidents. This point is an important one. In a general-equilibrium context, in which alternative risk-reduction activities must compete for scarce resources, asymmetric conservatism may reduce overall safety.[27]

Some critics have argued that current risk-assessment procedures understate risks (and hence the benefits of regulation) because so few substances have been tested adequately and because many assessments are restricted to a single health effect—cancer.[28] These are real problems that deserve attention, but they can not be solved by overestimating the risk of cancer from identified carcinogens. Indeed, such overestimates are likely to exacerbate the problem by diverting scarce resources away from dealing with a broader array of substance and health effects.

One could also argue that most people view cancer as a worse way of dying than other diseases (such as heart attacks), so that carcinogenic risk assessment should be especially conservative. The appropriate way of reflecting that preference, however, is to assign greater weight to cancer deaths than to other fatalities in making risk-management decisions, not to differentially overstate cancer risks.[29]

### *Lead vs. Benzene*

The debate about EPA's most recent rules limiting lead in gasoline illustrates some of the distortions that can arise from asymmetric conservatism in assessing carcinogenic and noncarcinogenic risks. Exposure to high levels of lead can lead to neurological damage in children.[30] Extensive statistical analyses have shown that decreasing lead in gasoline would greatly reduce the number of children exposed to that risk. EPA estimates that its new, more stringent rule will reduce by more than 150,000 per year the number of children with blood lead measurements above the level set by the Centers for Disease Control (CDC) for follow-up care. Recently completed studies also suggest that exposure to lead raises blood pressure in adult males, and that the rule could eliminate each year more than 1.5 million cases of hypertension and more than 5000 deaths from heart attacks and other diseases related to blood pressure.

Because the rule will increase the cost of manufacturing leaded gasoline, it is expected to reduce misfueling — the misuse of leaded gasoline in cars

that need unleaded to protect their pollution control devices. Thus emissions of hydrocarbons, nitrogen oxides, and carbon monoxide also should decrease. Finally, because lead and the scavengers added to leaded gasoline combine to corrode exhaust system and engine parts, reducing lead in gasoline should reduce automobile maintenance expenditures by almost $1 billion per year.

On the negative side, the refining processes used to meet octane requirements with less lead are likely to increase the amounts of aromatics in gasoline, including benzene, a leukemogen that EPA has listed as a hazardous air pollutant. Some observers saw this increase in the benzene content of gasoline as a potentially serious adverse side effect of the proposal to reduce lead in gasoline.[31]

Any comparisons of the benzene and lead risks are distorted, however, by the very different degrees of conservatism employed in their estimation. The benzene risk estimate followed the conservative procedures outlined above for carcinogens. Based on studies of workers exposed to several hundred parts per *million* of benzene, the estimation extrapolated linearly down to current ambient concentrations of several parts per *billion*. A variety of other conservative assumptions also biased the CAG unit-risk estimate upward (Nichols, 1983).

The lead analysis, in contrast, was based on data on exposures at the relevant levels. Moreover, the quantitative risk estimate for children did not include any effects below the blood-lead levels defined as toxic by the CDC, though many studies suggest that damage can occur at lower levels. (These lower levels in dispute are roughly one-half the levels at which serious effects are well accepted, not hundreds or thousands of times lower, as with benzene.)

The estimates of the effects of lead on adult health are even more restricted. Because the study linking blood lead and blood pressure showed a statistically significant result only for males, women were not included in the estimate. (These results were consistent with animal studies, in which lead induced elevated blood pressure in male, but not female, rats.) And because the best data on cardiovascular risk factors were for white males aged 40–59, no estimates were included for nonwhites or for whites outside that age range. Had the procedures employed been analogous to those used with carcinogens, the results for white males aged 40–59 would have been extended to women, nonwhites, and other age groups, yielding estimates many times higher. (The benzene estimates were based on men of working age, but were applied to both sexes and all ages. When animal tests are used for extrapolation with carcinogens, the results from the sex–species group with the highest excess risk are applied to people of all ages and both sexes.)



Because of these differences, any direct comparison of the lead and benzene risks as traditionally estimated by EPA is misleading and might have led to undercontrol of lead to avoid the exaggerated risks estimated for benzene. Two factors prevented this outcome. First, even with the conservative approach used for benzene, the estimated risks are relatively small; the estimate for evaporative emissions from gasoline marketing (service stations plus wholesale terminals) is less than four cases of leukemia per year from benzene.[32] In contrast, several thousand deaths may be caused by blood-pressure-related diseases associated with lead, and hundreds of thousands of children are at risk of neurological problems and other health effects. Second, and probably more important politically, additional analysis showed that reducing lead in gasoline will reduce net benzene emissions. As the EPA rule makes leaded gas more expensive, it will reduce the misfueling problem; as fewer pollution control catalysts are disabled, emissions of benzene and other hydrocarbons will decline.[33]

We cannot expect that the bias introduced by asymmetric conservatism will always be so fortuitously overcome. This example illustrates what we believe to be a broader problem: EPA spends more of its marginal resources regulating carcinogens, as compared to other types of environmental hazards, than is warranted by their likely benefits for public health and welfare. The reasons for this misplaced emphasis include failures to apply rational principles of risk management, public misperceptions about relative risks from carcinogens and noncarcinogens, and statutes that reflect these misperceptions. Asymmetric conservatism, which overstates carcinogen risks, also contributes to the problem.

### *Misperceptions of Probabilities*

Public policy with respect to risks also is complicated by the great difficulty that individuals have in perceiving small risks accurately and in processing such information. An extensive literature has developed on this subject, showing through experiments and other means the very serious biases that pervade perceptions of different risks. Small changes in the ways in which choices involving uncertainty are presented, for example, can lead to radically different choices. [Kahneman et al. (1982) provide a useful summary of much of this research.]

These perceptual problems raise serious philosophical issues about the criteria that public decision makers should use in questions involving risk. In particular, what relative weight should be given to perceived risks, which may be based on very poor information and are subject to various biases, as compared to estimates derived from experts, and when should either class of estimates be recalibrated to account for known psychological pro-

pensities? This subject goes beyond the scope of this paper. We believe that whatever position one takes on this question, however, it is not appropriate to distort the "objective" assessment to make it correspond with public perceptions.

## IV. THE EXPECTED-VALUE APPROACH

The major conceptual alternative to conservative analysis is to compute the expected value of risk in terms of a probability of loss. Under that approach, the risk assessors make alternative estimates and compute the weighted average, weighting each estimate by the probability that it is correct. If one model predicts a risk of $10^{-4}$ at the relevant dose, while another predicts a risk of $10^{-3}$, and the assessors believe that the first estimate is three times more likely to be correct than the second, then the expected value is $0.75(10^{-4}) + 0.25(10^{-3}) = 3.25 \times 10^{-4}$.

The expected-value approach does not entail simple acceptance of the "most likely" estimate, which in the example above would be $10^{-4}$. Rather, it uses the full range of estimates. Thus it does not ignore asymmetries in the uncertainties, nor does it reject "high" estimates. This is important, because much of the expected gain from regulation often comes from a small probability that a substance poses a relatively high risk.[34]

### A. Risk Aversion

Most people, if given a choice, will not accept actuarially fair gambles that involve substantial sums of money. Suppose a lottery offers participants a 50–50 chance of doubling their annual income or losing it altogether. Although the expected value of that lottery is zero — not negative — most people would pay a substantial amount of money not to have to enter it. Decision analysts describe such individuals as risk-averse.

What about risks to health? Suppose chemical A poses a risk of cancer of $10^{-5}$ for each year of exposure. The risk from chemical B is uncertain; there is a 90 percent chance that it is "safe" (0 risk) and a 10 percent chance that it poses a risk of $10^{-4}$ (that is, 10 times higher than A). If they must be exposed to one of these substances, which one should risk-averse individuals choose? Intuition may suggest that they should prefer A, because with B they run the chance of exposing themselves to a much higher risk. In fact, however, they should be indifferent. The expected risk with B is $0.9(0) + 0.1(10^{-4}) = 10^{-5}$, precisely the same risk posed by A. And unlike the person offered a double-or-nothing lottery on her/his salary, these individuals are not offered a choice between a certain and an uncertain future. Rather they must choose between two risks involving the same stakes. Risk aversion applies to states, not probabilities. It is irrelevant to this choice, because what is at issue is not the potential outcome



("cancer" in both cases), but rather the probability of that event. Individuals, whether risk-averse or risk-neutral, should rely on unbiased probability estimates, computed as the expected value of the probability distribution on a probability.[35]

#### *Risk Aversion on Risks to Society*

This logic indicates that individuals should base their decisions on the expected value of the risk estimate, not a conservative estimate. Some observers argue that this conclusion should not be extended to societal decisions, where the risks will apply to many individuals. Continuing the example above, suppose the decision is being made by EPA, and that 1 million people will be exposed to the substance. If A is chosen, the decision maker can anticipate that about $10^{-5}(10^6) = 10$ extra cancers will occur each year. If B is chosen, however, there is a 10 percent chance that about $10^{-4}(10^6) = 100$ extra cancers will occur, and a 90 percent chance of none. Thus A offers 10 cases with relative certainty, while B offers a lottery with a 0.1 chance of 100 and a 0.9 probability of 0, with an expected value of 10 cases. Although both substances offer the same expected risk, some believe that society should be risk-averse and choose A. In the context of nuclear power, the analogous argument is that low-probability catastrophes involving thousands of deaths should receive more weight than the expected-value computation would assign them. The simultaneous death of 1000 people in the same town is seen as worse than the isolated deaths of 1000 otherwise identical people killed in individual incidents.[36]

We believe that, in general, society should be approximately risk-neutral in the type of case outlined above, that A and B should be treated as equivalent for regulatory purposes. (This leaves aside the possibility of gathering more information to resolve the uncertainties associated with B; if information can be obtained, it may make sense to do so. Indeed, if further use will provide such information, it will be preferable to allow the use of B.) The analogy to large-scale catastrophes is misleading; in the case of environmental carcinogens, the cancers are likely to be scattered widely in time and space. Moreover, the numbers involved, even if the "high" estimates are correct, are small relative to virtually any relevant measure, whether the overall risk of death or the risk of cancer in particular. Each year, for example, about 850,000 Americans contract cancer (U.S. Environmental Protection Agency, 1985a, p. vi). The number of cases involved in any given regulation is a tiny fraction of that total, even when "conservative" (i.e., high) estimates are used.

A recent EPA study of airborne toxic substances, for example, employing all of the standard conservative assumptions, estimates the total cancer risk for several dozen substances to be less than 2000 cancer cases a year,

or about 0.2 percent of all cancers (U.S. Environmental Protection Agency, 1985a, p. vi). Moreover, many of these cases are attributed to "products of incomplete combustion," a catchall category including hundreds of types of sources and millions of individual sources (cars, wood stoves, and virtually every other activity that involves the burning of fuel). Thus most individual regulations are likely to involve a handful of cases. Given such figures, it is difficult to argue that risk assessments for carcinogens should be conservative because of risk aversion with respect to "catastrophic" outcomes, although we also recognize that political prominence and actual risks are rarely related linearly (and at times not even monotonically).

### B. Erring on the Side of Safety

We have used "risk aversion" in its traditional decision-analytic sense, to describe the common finding that most people will not accept fair gambles involving large stakes. When people use the term in the context of risks to health, however, they usually mean that they place a large value on avoiding or reducing such risks. Risk aversion in this latter sense—quite different from that used by decision analysts—is probably the primary motivation behind conservative risk-assessment procedures. As we have shown, however, such procedures can lead to less rather than more protection.

This second, less technical definition of risk aversion is perfectly compatible with an expected-value approach to risk assessment. If society wants to place a high value on risk reduction, it should do so through the decision rules for risk management, not by distorting risk-assessment procedures. As economists, we naturally think in terms of marginal valuations; the desire to give extra weight to safety can be expressed by attaching a high value or shadow price to risk reduction. In contrast to conservative risk assessment, such an approach offers three major advantages. It leads to greater consistency across regulations and to an emphasis on tackling the most serious problems—rather than those with the most conservative assessments—first. Second, it makes the value trade-offs more explicit, thus facilitating public debate and making politically responsible risk managers more accountable. Third, it avoids the long-term consequences of always crying wolf through rampant conservatism. As decision makers become more aware of the conservatism in risk assessments, some are likely to compensate by making ad hoc, downward mental adjustments to the formal estimates presented to them.

### C. Toward More Realistic Risk Assessments

Ideally, risk assessors would report the full range of possible estimates and their associated probabilities. The range and the probabilities would



reflect the uncertainties present in all stages of the assessment process. Benefit–cost estimates and other analyses also would carry through this probabilistic assessment. We would urge that risk managers then rely on the expected value of risk in making their choices. Even if they chose to be conservative (e.g., relying on the 90th percentile of the risk estimate), at least the choice would be explicit and debates about the procedure would not be handicapped by the illusion that it was based on science. The line between assessment and management would be sharp, with the responsibility for policy judgments about margins of safety and the like clearly lodged with politically accountable policy makers, not hidden in a series of obscure, "prudent" decisions by scientists.

Unfortunately, current practice is far from this ideal, nor is it likely to be achieved in the foreseeable future. The obstacles are both perceptual and substantive.

Many will perceive efforts to reduce conservatism in risk assessment as an attempt to politicize risk assessment and to reduce the protection afforded the public. Such suspicions are reinforced by the fact that many opponents of conservative risk-assessment procedures have strong financial or ideological interests in reducing regulation. Thus it is critical that efforts to reduce the use of "worst-case" procedures be presented (accurately) as an attempt to remove "political" judgments from risk assessment, and that the problem be placed in a general framework that highlights the inefficiencies caused by conservatism in allocating resources across different types of risks.

The major substantive difficulty lies in assessing the probabilities that alternative assumptions and models are correct. In a few cases (e.g., uncertainties due to the limited numbers of animals employed in experiments), probabilities can be estimated using standard statistical techniques. The major sources of uncertainty, however, are less tractable, because they reflect limited knowledge of the underlying biological and physical processes involved. These problems are particularly acute in the case of dose–response models. The obvious answer is to obtain subjective probability estimates from the relevant experts. Unfortunately, the experts are likely to resist assigning probabilities and, inevitably, they will disagree. Designing a system for eliciting and combining subjective estimates from experts that achieves widespread acceptance is a formidable task.[37]

Despite these daunting obstacles, we believe that progress is possible. The first step is for assessors to report a broader range of estimates based on alternative "plausible" assumptions, not just the "plausible upper bound." The problem, of course, is that the resulting range will be huge, leaving risk managers with a choice between continuing to rely on the upper bound or making scientific judgments for which they have little or no training, thus intruding into the risk assessment process. We strongly

suspect that most will take the first option.[38] Nonetheless, routine reporting of such ranges could play a useful role in educating decision makers and the public about the extraordinary uncertainties involved and the degree to which current procedures may bias estimates upward.

A second useful step would be to tackle some of the more obvious and solvable sources of conservatism. One such source, already under review by EPA, is the method used to estimate the parameters of the dose–response model from experimental data; as discussed earlier, simply switching from the current method of using the 95 percent upper bound estimate of the linear term in the multistage model to an estimator based on central tendency could reduce the conservatism of carcinogenic risk estimates by about a factor of 2. Exposure assessment offers similar opportunities for reducing conservatism; estimates could be based on more realistic assumptions about dispersion and activity patterns. A slightly more ambitious task would be to develop alternatives that rely on the full range of experiments available, rather than simply the one yielding the highest risk estimate. Although each of these changes might yield only a "small" change (e.g., reducing the risk estimate by a factor of 2), the cumulative effect would be substantial.

The most ambitious step would be to develop subjective probability estimates from experts for the more fundamental uncertainties, in particular about the shape of the dose–response function. The prospects for the success of such an effort will improve as scientists' understanding of the fundamental processes increases and a greater consensus begins to emerge. In the meantime, our inability to achieve the ideal should not dissuade us from addressing the more obvious and tractable sources of conservatism in risk assessment. Until we do, we will go on misallocating the scarce resources available to reduce risk and to improve our general well-being.

## ACKNOWLEDGMENTS

The authors' work on this paper was supported in part by the Energy and Environmental Policy Center and the Business and Government Center of the John F. Kennedy School of Government, Harvard University.

## NOTES

1. Ruckelshaus's (1983) speech to the National Academy of Sciences (NAS) deals with the assessment–management distinction, focusing on assessment. Some of the steps taken to improve assessment and management at EPA are laid out in a small booklet issued shortly before he left office (U.S. Environmental Protection Agency, 1984b).

2. There are many alternative ways of defining the steps in risk assessment. EPA officials, for example, frequently describe it as a four-step process (U.S. Environmental Protection Agency, 1984b): (1) hazard identification (the qualitative determination that a substance poses some risk to human health); (2) dose–response assessment; (3) exposure assessment; and (4) risk characterization (which combines the results of the other steps to report risk estimates, uncertainties, and so on).

3. This simplified formulation implicitly assumes that regulations are directed at emissions, although our discussion of conservatism also applies to other types of regulations that might be employed. For example, restrictions may be placed on the times and places at which an activity can take place, thus reducing exposure without necessarily affecting the volume of emissions. See Nichols (1984, esp. Chap. 5) for a discussion of the implications of alternative regulatory targets.

4. In that case, the initial estimate, used to rank source categories for regulatory action, was 15.8 million kilograms (kg) of benzene per year from maleic anhydride plants. Better information on existing controls later lowered the estimate to 5.8 million kg, assuming full-capacity operation; using actual capacity utilization would have lowered the estimate to 3.3 million kg. After the rule was proposed, several plants closed or switched feedstocks—because the price of benzene, a petroleum derivative, sharply increased in the late 1970s—lowering the estimate of baseline emissions to less than 1 million kg per year (Nichols, 1983).

5. See Haigh et al. (1984) for a discussion of the uncertainties involved in estimating coke oven emissions and the reasons for the wide range in EPA's estimates.

6. Virtually all of the literature assumes that compliance will be perfect, i.e., that all plants will just meet the standard. Viscusi and Zeckhauser (1979) analyze a situation in which tightening the standard may reduce compliance (assuming fixed levels of enforcement), and show that the optimal standard in such cases will be less stringent than it would be with perfect compliance. Alternative assumptions, however, could lead to the opposite conclusion, that imperfect compliance increases the optimal stringency of regulations.

7. In developing regulations to ban certain wastes from landfills, for example, EPA is assessing the concentrations that the wastes might reach in groundwater once they leak out of a landfill. A key parameter in that assessment is the distance from the leak at which concentrations are estimated; the shorter the distance the higher the estimated concentration, and thus the more likely a ban. A survey of existing landfills has found that the distance between a landfill and the nearest well varies widely. EPA's Office of Solid Waste has proposed that the analysis used for regulatory purposes employ the minimum distance found in that survey; i.e., the analysis would assume the "worst-case" distance for all substances and locations. That assumption will, on average, substantially overestimate the concentrations that would occur in drinking water (personal communication from Dr. Ann Fisher, Office of Policy Analysis, EPA, Washington, D.C., September 11, 1985).

8. This assumes that the assessment is made prospectively, before the event occurs (as with potential contamination of groundwater). If the assessment is made after the contamination has occurred, damage estimates based on observed exposure may be understated because they ignore costs of averting behavior to avoid exposure. For a more complete discussion of the issues raised by averting behavior, see Zeckhauser and Fisher (1976).

9. An NAS committee on pesticide regulation, for example, concluded that extrapolation from animals to human beings was so tenuous that quantitative risk estimates should not be attempted in the absence of human data. The committee did not, however, reject the use of animal data for qualitative purposes or for setting priorities (National Academy of Sciences, 1980).

10. Surface area is assumed to be proportional to two-thirds the power of weight. Thus the two methods differ in their potency estimates by a factor equal to the ratio of the weight of a person to the weight of the animal, raised to the one-third power.

11. This assumes that we cannot identify who is more sensitive and who is less sensitive. From an ex ante perspective, everyone runs the same risk, the weighted average of the risks of all of the different (anonymous) groups. If we can identify the sensitive groups, however, variations in sensitivity may matter for two reasons. If those who are more sensitive can take cost-effective action to protect themselves, the optimal (efficient) level of exposure is likely

to be higher than if everyone faces the same risk (or we cannot tell who is more sensitive). On the other hand, if the variations in sensitivity are large, so that some identifiable groups face unusually large risks, equity may argue for a tighter level of control, as may politics because identified losers speak loudly.

12. Our discussion of the procedure for estimating the parameters of the multistage model is based primarily on a series of meetings that one of the authors (Nichols) attended at EPA during 1984 and 1985. At those meetings, Dr. Todd Thorslund of the CAG provided a concise discussion of the reasoning behind the use of the 95 percent confidence limit, and also presented several useful examples, based on Monte Carlo simulations, of the sensitivity of the MLE to small changes in experimental data.

13. When it uses the pure linear model with human epidemiological data, the CAG typically does not use an upper-confidence estimate of the linear term, because that model inherently forces linearity.

14. In 1984, EPA's Risk Assessment Forum took up the issue of the MLE versus the upper confidence approach to estimating the parameters of the multistage model. The Forum recommended additional review of an "average likelihood estimator" developed by Dr. Todd Thorslund of the CAG. That estimator is based on a weighted average of the linear term, with the relative weights equal to the associated value of the likelihood function. That method had several desirable properties, but lacked a firm theoretical basis. The CAG is now investigating another approach, which uses the underlying theory of the multistage model to place restrictions on the relative values of the parameters and assures that the maximum likelihood estimate has a positive linear term (personal communication from Dr. Robert McGaughy, Office of Research and Development, EPA, Washington, D.C., September 16, 1985).

15. At low doses, risk is roughly proportional to dose squared under the two-hit model; thus lowering the dose by a factor of 1000 reduces risk by a factor of about $1000^2 = 1$ million.

16. This problem is well known. See Nichols (1983) for an illustrative example showing the close correspondence of several of the models in the observable range and their wide divergence at low doses.

17. We argue later in the text that risk estimates should be based on expected values. In the case of the dose–response model, that would require making risk estimates with alternative models, and then weighting the estimates by experts' subjective probabilities that each model is correct. Because the nonlinear models generally predict low-dose risks that are so much smaller than the linear model's estimate, the resulting expected value would be roughly equal to the linear model's predicted risk times the probability that the linear model is correct. Note that this expected dose–response function would be roughly linear at low doses, but its slope would be smaller than that based on the pure linear model; e.g., if experts estimated that the linear model had a 50 percent chance of being correct, the slope (and estimated low-dose risk) would be one-half the value estimated with the linear model.

18. See Maugh (1978) for a brief, nontechnical summary of the debate about low-dose risks from carcinogens.

19. If the risk levels are high enough, supralinear low-dose functions also could arise if individual dose–response functions were linear, but the slopes of those function varied across the population. The supralinearity can arise because an individual cannot die twice, or if a strong selection effect is exerted so the population surviving at higher doses is significantly less susceptible to the risk. [See Shepard and Zeckhauser (1980) for a discussion of these issues.] For the types of risks discussed here, however, the probabilities of death are tiny relative to the magnitudes needed to make either of these effects more than a theoretical nicety.

20. Campbell et al. (1982) calculated an expected value of the risk estimate based on subjective estimates of the probabilities that the alternative assumptions were correct. Their subjective estimates, however, were illustrative and were not based on any systematic effort to poll scientists.



21. The lognormal is particularly convenient because the logs of the parameters are normally distributed. Note that if the original variables are to be multiplied, we can add the logs of the variables. Let s be the standard deviation of the log of each variable. Since the log of the 90th percentile is 1.282s from the mean of the log of the variable, the sum of the logs of the five 90th percentile estimates will be $5(1.282) = 6.41$s from the sum of the means of the logs of those variables. Assuming independence, the variance of the sum of the logs of the variables will be simply the sum of the individual variances, or $S^2 = 5s^2$, and the standard deviation will be $S = 5^{0.5}$s. Thus the sum of the logs of the five 90th percentile estimates will lie $6.41s/(5^{0.5}s) = 2.87$ standard deviations from the sum of means of the logs; from a table of values of the cumulative normal distribution, we find that 2.87 standard deviations lies at the 99.8th percentile.

22. EPA's proposed cancer assessment guidelines (49 *Federal Register* 46294, 1984) contain a provision that may help in this regard. Under the proposed guidelines, the risk-assessment process would produce a formal rating of the "strength of the evidence" that a substance is a carcinogen, based on the type of studies available (e.g., human epidemiology versus carcinogenicity in multiple animal species vs. carcinogenicity in a single species). These ratings would then be reported along with the quantitative estimates. At present, for purposes of quantitative risk assessment, strength of evidence is used as a binary variable; quantitative assessments are prepared for those substances that pass some threshold, whether by a small margin or a large one. Conversely, those substances that fail to pass the threshold are not assessed, so implicitly their risks are treated as zero. For those substances that pass, a verbal description of the strength of the evidence is included in assessments, but no single rating is given, and summary charts typically do not contain the description. Although a step in the right direction, these ratings will not cover other differences in conservatism across assessments, nor will they provide a quantitative basis for adjusting risk estimates on the basis of the likelihood that a substance is truly a human carcinogen. Development of a probability-weighting scheme also would facilitate the assessment of substances for which the evidence is now considered too limited to justify a quantitative assessment.

23. If a linear and a convex function are both fit through the origin and a common point, the convex function generates lower estimates of risk at every dose below the common point; over a substantial portion of that range, however, the convex function has a larger slope (greater marginal risk). A model with risk proportional to dose-squared, for example, generates higher estimates of marginal risk than the linear model for all doses greater than half the dose at the common point to which both models are fit.

24. Under the Clean Air Act, for example, EPA is directed to set ambient standards that provide an "adequate margin of safety;" these standards are based primarily on epidemiological or laboratory studies of humans exposed in the relevant range.

25. On air bags, for example, see Huelke and O'Day (1981). They compare their estimate of the lives that would be saved by air bags or other passive restraints with those made by the National Highway Traffic Safety Administration (NHTSA), and conclude that the latter are "much higher." In no case, however, are the NHTSA estimates more than double those of Huelke and O'Day; in contrast, estimates of the risks from environmental carcinogens routinely vary by orders of magnitude.

26. In discussing the drawbacks of conservative assessments, Lave (1981, p. 35) notes: "An unintended result is to exaggerate effects most in those cases where there is the greatest ignorance. Thus this procedure allocates the most resources to the cases where there is the most uncertainty."

27. Consider the following example. Perceived safety is a function of expenditures on reducing two risks, $x_1$ and $x_2$:

$$S = (1 - e^{-0.01x_1})\,(1 - ge^{-.01x_2}),$$

where g is the factor by which the second risk is overestimated. If g = 1, then received safety is the same as actual safety. Disposable income is $1000 - x_1 - x_2$, and utility is given by

$$U(Y,S) = (1 - e^{-0.001Y})S.$$

The optimum is $x_1 = x_2 = 215.8$, $Y = 568.4$, and $S = 0.781$. If the second risk is overestimated by a factor of 100 (that is, g = 100), expenditures on safety rise sharply, dropping disposable income to Y = 261.7, but the overall level of safety falls to S = 0.748 because an extremely inefficient mix of safety expenditures is chosen. Asymmetric conservatism leads to excessive expenditures on the second risk, and the income effect reduces expenditures on the first. Moreover, with the same overall level of expenditure on safety, a much higher level of overall safety, S = 0.951, could be achieved if unbiased risk estimates were used.

28. Controlling carcinogens can reduce noncancer risks for at least two reasons: (1) A carcinogen may have other adverse effects; e.g., acrylonitrile is a carcinogen and also contributes to the formation of ozone. (2) Controls may reduce emissions of more than one pollutant; e.g., vapor controls to reduce evaporative emissions of benzene from gasoline also capture other gasoline components, some of which also may be carcinogenic and most of which contribute to ozone. Thus, it is important that analyses of the costs and benefits of controlling carcinogens include noncancer benefits. The relationship between cancer-related and other benefits, however, varies widely across substances and source categories, so that the cancer benefits are a very poor proxy for total benefits.

29. Consider other cases where competing policies yield outputs that are similar but which are valued differently. For example, suppose that two reading programs, A and B, are being evaluated, where A is more effective with poorer students and B is more effective with average students. If we value reading gains by poorer students more highly than those by average students, we should not account for this preference by overestimating the effects of program A. Rather, we should give greater weight to gains at the bottom of the scale than to those in the middle.

30. The estimates of the effects of reducing lead in gasoline in the following discussion all are drawn from the regulatory impact analysis (U.S. Environmental Protection Agency, 1985b).

31. Perhaps not surprisingly, these concerns were raised most forcefully by representatives of the Lead Industries Association. One of the authors (Nichols), however, also was present at several meetings at which EPA officials raised the issue as a potentially serious problem, and requested that additional work be done to resolve the question. When the final rule was announced at a press conference, a reporter for a national wire service asked if reducing lead would not also increase cancer.

32. EPA estimates 123 cases of leukemia due to benzene exposure from gasoline marketing over a 35-year period (U.S. Environmental Protection Agency, 1984a, Table 6-9), or about 3.5 cases per year.

33. EPA estimates that although the reduction in lead in gasoline will result in a small increase in evaporative emissions of benzene from service stations and other parts of the gasoline marketing system, that increase will be swamped by reductions in automobile tailpipe emissions of benzene, with a net reduction of about 18,000 tons per year (U.S. Environmental Protection Agency, 1985b, Table VI-12).

34. See Page (1978) for further discussion of the implications of asymmetries in the losses associated with false positives and false negatives in regulatory decisions.

35. The theory of risk aversion was developed by calibrating complementary probabilities for two unchanging reference prizes to be equivalent to a certainty of receiving the prize being assessed (Raiffa, 1968). To be linear in utilities, the purpose of the exercise, requires that one be linear in the the probability of winning the better reference prize.

36. Wilson (1974), for example, suggests that the loss may rise with the square of the number of fatalities; i.e., an accident killing 1000 people simultaneously is as bad as 1 million individual deaths scattered in time or space. Slesin and Ferreira (1978) argue for an even more sharply nonlinear function, with damage rising with the cube of the number of fatalities; i.e., a 1000-person accident is equivalent to 1 billion individual deaths. As Schelling (1968) points out, however, damage may rise less than proportionately to the number of deaths because a significant part of the damage caused by a death is the bereavement of close relatives and friends; in concentrated accidents many of those who would feel the greatest bereavement die simultaneously. We find this argument compelling, and believe that in most cases society should not be "risk-averse" with regard to accident sizes. As discussed in the text, however, one's views on this issue are of little relevance to the question of environmental carcinogens.



37. See Raiffa and Zeckhauser (1981) for a discussion of some of the issues involved in eliciting such opinions from experts.

38. CAG, for example, notes that the "plausible lower bound" in all of its low-dose cancer estimates is zero, so in some sense current estimates already include a range, though not a very informative one.

## REFERENCES


Anderson, E. L., "Quantitative Approaches in Use in the United States to Assess Cancer Risk." Paper presented at Workshop on Quantitative Estimation of Risk to Human Health from Chemicals, Rome, July 31 (1982).

Campbell, G. L., D. Cohan and D. W. North, *The Application of Decision Analysis to Toxic Substances: Proposed Methodology and Two Case Studies*. Washington, DC: Office of Pesticides and Toxic Substances, U.S. Environmental Protection Agency, Contract 68-01-6054, 1982.

Crandall, R. W. and L. B. Lave, *The Scientific Basis of Health and Safety Regulation*. Washington, DC: Brookings Institution, 1981.

Crouch, E. and R. Wilson, "Interspecies Comparison of Carcinogenic Potency." *Journal of Toxicology and Environmental Health* 5(1979):1095–1118.

Crouch, E. and R. Wilson, "Regulation of Carcinogens." *Risk Analysis* 1(1981):47–57.

Galluzzo, N. B. and D. Glassman, "Recalculated Benzene Emissions, Population Exposures, and Risks." In *CMA Presentation at the EPA Hearing on the Proposed NESHAP for Benzene Emissions from Maleic Anhydride Plants*. Washington, DC: Chemical Manufacturers' Association, August 21 (1980).

Haigh, J. A., D. Harrison and A. L. Nichols, "Benefit-Cost Analysis of Environmental Regulation: Case Studies of Hazardous Air Pollutants." *Harvard Environmental Law Review* 8(1984):395–434.

Hoel, D. G., N. L. Kaplan, and M. W. Anderson, "Implication of Nonlinear Kinetics on Risk Estimation in Carcinogens." *Science* 219(1983):1032–1037.

Huelke, D. F. and J. O'Day, "Passive Restraints: A Scientist's View." In R. W. Crandall and L. B. Lave (eds.), *The Scientific Basis of Health and Safety Regulation*. Washington, DC: Brookings, Institution, 1981.

Kahneman, D., P. Slovic, and A. Tversky, (eds.), *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press, 1982.

Lave, L. B., *The Strategy of Social Regulation: Frameworks for Policy*. Washington, DC: Brookings Institution, 1981.

Lave, L. B. (ed.), *Quantitative Risk Assessment in Regulation*. Washington, DC: Brookings Institution, 1982.

Maugh, T. H., "Chemical Carcinogens: How Dangerous Are Low Doses?" *Science* 202(1978):37–41.

Miller,C., *Exposure Assessment Modeling: A State-of-the-Art Review* Washington, DC: U.S. Environmental Protection Agency, EPA-600/3-78-065, 1978.

National Academy of Sciences, *Regulating Pesticides*. Washington, DC: Report of the Committee on Prototype Explicit Analyses for Pesticides, 1980.

National Academy of Sciences, *Risk Assessment in the Federal Government: Managing the Process*. Washington, DC: Report of Committee on the Institutional Means for Assessment of Risks to Public Health, 1983.

Nichols, A. L., "The Regulation of Airborne Benzene." In T. C. Schelling (ed.), *Incentives for Environmental Protection*. Cambridge, MA: MIT Press, 1983.

Nichols, A. L., *Targeting Economic Incentives for Environmental Protection*. Cambridge, MA: MIT Press, 1984.

Page, T., "A Generic View of Toxic Chemicals and Similar Risks." Pasadena, CA: California Institute of Technology, Social Science Working Paper, January, (1978).

Raiffa, H., *Decision Analysis*. Reading, MA: Addison-Wesley, 1968.

Raiffa, H., W. Schwartz, and M. Weinstein, "Evaluating Health Effects of Societal Decisions and Programs." In *Decision Making in the Environmental Protection Agency*. Washington, DC: National Academy of Sciences, 1977.

Raiffa, H., and R. J. Zeckhauser, "Reporting of Uncertainties in Risk Assessment." Cambridge, MA: unpublished paper, Harvard University, 1981.

Ruckelhaus, W. D., "Science, Risk and Public Policy." Speech before the National Academy of Sciences, Washington, DC, June 22 (1983).

Schelling, T. C., "The Life You Save May Be Your Own." In S. Chase (ed.), *Problems in Public Expenditure Analysis*. Washington, DC: Brookings Institute, 1968.

Schneiderman, M. A., et al., "From Mouse to Man." *Annals of the New York Academy of Sciences* 246(1975):237–248.

Shepard, D. and R. J. Zeckhauser, "Long-Term Effects of Interventions to Improve Survival in Mixed Populations." *Journal of Chronic Disease* (1980):413–433.

Slesin, L. and J. Ferreira, "Social Values and Public Safety: Implied Preferences Between Accident Frequency and Severity." Cambridge, MA: Massachusetts Institute of Technology, September (1978).

Spengler, J. D. and K. Sexton, "Indoor Air Pollution: A Public Health Perspective." *Science* 221(1983):9–17.

U.S. Environmental Protection Agency, *Benzene Emissions from Maleic Anhydride Industry: Background Information for Proposed Standards*. Research Triangle Park, NC: Emissions Standards and Engineering Division, February, EPA-450/3-80-001a, 1980.

U.S. Environmental Protection Agency, *Evaluation of Air Pollution Regulatory Strategies for Gasoline Marketing Industry*. Washington, DC: Office of Air and Radiation, July, EPA-450/3-80-012a, 1984a.

U.S. Environmental Protection Agency, *Risk Assessment and Risk Management: Framework for Decision Making*. Washington, DC: U.S. Environmental Protection Agency, December, EPA 600/9-85-002, 1984b.

U.S. Environmental Protection Agency, *The Air Toxics Problem in the United States: An Analysis of Cancer Risks for Selected Pollutants*. Washington, DC: Office of Air and Radiation and Office of Policy, Planning and Evaluation, 1985a.

U.S. Environmental Protection Agency, *Costs and Benefits of Reducing Lead in Gasoline: Final Regulatory Impact Analysis*. Washington, DC: Economic Analysis Division, February, EPA 230-05-85-006, 1985b.

Viscusi, W. K. and R. J. Zeckhauser "Optimal Standards with Incomplete Enforcement." *Public Policy* 27(1979):437–456.

Wilson, R., "Examples in Risk–Benefit Analysis." Paper presented at Conference on Advanced Energy Systems, Denver, June (1974).

Zeckhauser, R. J. and A. Fisher, "Averting Behavior and External Diseconomies." Discussion paper 41D, John F. Kennedy School of Government, Harvard University, April (1976).