Good Science, Bad Regulation, and Toxic Risk Assessment

Howard Latin†

Regulation of toxic substances is an extremely complex, uncertain, and controversial enterprise. The regulatory process is customarily divided into two discrete functions: *risk assessment* ostensibly is a scientific activity that develops estimates of health hazards at varying exposure levels, while *risk management* is a political activity that balances competing interests and values to determine whether identified toxic risks should be considered unacceptable or tolerable.[1] This sharp distinction between the scientific and social policy dimensions of toxics regulation is embodied in the Environmental Protection Agency's (EPA's) guidelines for estimating carcinogenic hazards, which provide that risk assessments must "use the most scientifically appropriate interpretation" and should "be carried out independently from considerations of the consequences of regulatory action."[2] The requirement for adoption of the "most scientifically appropriate interpretation" reflects EPA's current priority on attaining "good science" in risk-assessment proceedings.[3] In other words, EPA and other federal

† Professor, Rutgers University School of Law at Newark. J.D. 1974, Boalt Hall School of Law, University of California at Berkeley. I presented a preliminary version of this Article at a symposium entitled "Grappling With Risk Assessment: On the Frontier of Science and Law," held at Granlibakken, Lake Tahoe, California in October, 1986. I thank the organizers of the symposium, Dr. David Goldsmith, Associate Director, Toxic Substances Research and Teaching Program, University of California at Davis, and Professor John Dwyer, Boalt Hall School of Law, University of California at Berkeley, for their encouragement and support. I also thank Dr. Ellen Silbergeld of the Environmental Defense Fund and Professors Bruce Ackerman, John Dwyer, Don Elliott, John Leubsdorf, Jerry Mashaw, Frederica Perera, Jim Pope, Richard Stewart, and Peter Strauss for their comments on previous drafts of this Article.

1. *See, e.g.*, NATIONAL RESEARCH COUNCIL, RISK ASSESSMENT IN THE FEDERAL GOVERNMENT: MANAGING THE PROCESS 3 (1983) [hereinafter NATIONAL RESEARCH COUNCIL]:

> Risk assessment is the use of the factual base to define the health effects of exposure of individuals or populations to hazardous materials and situations. Risk management is the process of weighing policy alternatives and selecting the most appropriate regulatory action, integrating the results of risk assessment with engineering data and with social, economic, and political concerns to reach a decision. . . . At least some of the controversy surrounding regulatory actions has resulted from a blurring of the distinction between risk assessment policy and risk management policy.

See also Albert, *U.S. Environmental Protection Agency Revised Interim Guideline for the Health Assessment of Suspect Carcinogens*, 19 BANBURY REPORT, RISK QUANTITATION AND REGULATORY POLICY, 307, 308 (D. Hoel, R. Merrill & F. Perera eds. 1985).

2. EPA Guidelines for Carcinogen Risk Assessment, 51 Fed. Reg. 33,992, 33,992-93 (1986) [hereinafter EPA Carcinogen Guidelines].

3. The term "good science" in quotation marks is used to denote Agency attempts to base risk assessments only on the most plausible present scientific evidence, however provisional the scientific

agencies now stress the need for scientifically credible risk assessments and presume that their analyses should be grounded exclusively on the *best available* scientific theories and data even if the resulting predictions do not achieve the degree of reliability ordinarily required for valid scientific conclusions.

This Article challenges the conventional view that scientific perspectives should dominate the risk-assessment process. To paraphrase Talleyrand, risk assessment is too important and too uncertain to be left exclusively to the risk-assessors. I contend instead that social policy considerations must play as prominent a role in the choice of risk estimates as in the ultimate determination of which predicted risks should be deemed unacceptable.

Part I of this Article evaluates the risk-assessment principles in EPA's carcinogen guidelines. Part II examines risk-assessment practices adopted in EPA's proceedings on the carcinogen benzene and questions the scientific basis for the Agency's 1984 decision to control two categories of benzene sources while exempting other types of dischargers from regulation.[4] These discussions show that risk-assessment efforts in regulatory proceedings seldom achieve professionally accepted standards of scientific validity and inevitably entail implicit or explicit policy judgments.

Part III of the Article describes the social ramifications of EPA's current emphasis on "good science." In practice, this risk-assessment focus is likely to result in reduced public protection against potential toxic hazards, increased regulatory decisionmaking costs, and expanded opportunities for obstructive behavior by Agency bureaucrats or private parties hostile to toxics regulation. These consequences might be tenable if they were the product of an explicit political decision, but they should not arise unintentionally from the Agency's single-minded pursuit of "good science" in an area where reliable scientific conclusions are difficult or impossible to attain.

Part IV describes social policy criteria that agencies could use to supplement scientific evidence on toxic hazards. These factors include the hierarchy of legislative priorities in particular regulatory statutes, the expense and time requirements associated with individualized assessments

judgments may be.

4. *See* EPA National Emissions Standards for Hazardous Air Pollutants; Regulation of Benzene; Response to Public Comments, 49 Fed. Reg. 23,478, 23,480 (1984) [hereinafter EPA Regulation of Benzene]; EPA National Emission Standards for Hazardous Air Pollutants; Benzene Equipment Leaks (Fugitive Emission Sources), 40 C.F.R. §§61.110-.112 (1987); EPA National Emission Standards for Hazardous Air Pollutants; Proposed Standards for Benzene Emissions from Coke By-Product Recovery Plants, 49 Fed. Reg. 23,522 (1984) [hereinafter EPA Proposed Standards for Benzene Emissions]; EPA National Emission Standards for Hazardous Air Pollutants; Benzene Emissions from Maleic Anhydride Plants, Ethylbenzene/Styrene Plants, and Benzene Storage Vessels; Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558 (1984) [hereinafter EPA Withdrawal of Proposed Standards].

of recurring scientific issues, the potential for catastrophic results from erroneous risk estimates, and the likelihood that specific uncertainties can or cannot be resolved in the near future. Contrary to EPA's carcinogen guidelines, analysis of regulatory purposes and possible social consequences—not attempts at "good science" alone—should shape risk-assessment efforts. I believe statutory preferences for safety from toxic substances should not be undermined by low-visibility adoption of speculative risk-assessment practices that cannot be grounded on reliable science. Nevertheless, my objective in Part IV is not to advocate the policies I personally regard as desirable, but instead to examine a range of social policy criteria that could be incorporated in the risk-assessment process after appropriate public discussion.

Three interconnected themes, which are developed throughout the ensuing discussion, support the Article's central thesis that explicit social policy choices should influence agency selections of risk-assessment principles and specific risk estimates:

Inadequate scientific knowledge and inadequate data usually prevent derivation of risk estimates based on reliable science. Toxic risk assessment suffers from fundamental uncertainties about causal mechanisms for cancer and other hazards, extrapolative relationships between high-dose and low-dose responses and between animal test data and human risks, latent effects and latency periods, special sensitivities in exposed subpopulations, synergistic or co-carcinogenic effects of various substances, past and present exposure levels, dispersion patterns for contaminants, and virtually every other area of required knowledge.[5] These uncertainties generally preclude reliable assessments of relevant effects, and there is no scientific consensus on how they should be resolved. For example, conflicting risk estimates submitted in Food and Drug Administration (FDA) proceedings on saccharine varied by more than a millionfold;[6] and predictions

5. For example, William Ruckelshaus, then Administrator of EPA, observed:

> In assessing a suspected carcinogen . . . there are uncertainties at every point where an assumption must be made: in calculating exposure, in extrapolating from high doses where we have seen an effect to the low doses typical of environmental pollution; in what we may expect when humans are subjected to much lower doses of a substance that, when given in high doses, caused tumors in laboratory animals; and finally, in the very mechanisms by which we suppose the disease to work.

Ruckelshaus, *Science, Risk, and Public Policy*, 221 SCIENCE 1026, 1027 (1983). For policy-oriented discussions of the effects of scientific and medical uncertainties on toxic substances regulation, see NATIONAL RESEARCH COUNCIL, DECISION MAKING IN THE ENVIRONMENTAL PROTECTION AGENCY (1977) [herinafter EPA DECISION MAKING]; Latin, *The "Significance" of Toxic Health Risks: An Essay on Legal Decisionmaking Under Uncertainty*, 10 ECOLOGY L.Q. 339 (1982); McGarity, *Substantive and Procedural Discretion in Administrative Resolution of Science Policy Questions: Regulating Carcinogens in EPA and OSHA*, 67 GEO. L.J. 729 (1979).

6. *See* OSHA, Identification, Classification and Regulation of Potential Occupational Carcinogens, 45 Fed. Reg. 5002, 5200 (1980) (codified at 29 C.F.R. §§1900.101-1990.152 (1987)) [hereinaf-

of the hazards posed by TCE, a drinking-water contaminant, varied by many millions.[7] One discussion of TCE regulation noted that the "estimates provide a range of uncertainty equivalent to not knowing whether one has enough money to buy a cup of coffee or pay off the national debt."[8]

Under current regulatory practices, Agency scientists produce risk assessments that seldom approach the level of reliability normally expected of scientific findings; indeed, many estimates are little more than educated guesses.[9] Yet, the choice among competing estimates—a prediction of only a minuscule hazard or one a million times greater—can determine whether toxic exposures are characterized as "acceptable" or "unacceptable" irrespective of any values in the risk-management process. Absent a scientific consensus on which risk-assessment principles should be applied, I contend that an agency's choice among competing risk estimates should not be exclusively a result of provisional scientific judgments. If substantial uncertainty exists about the extent of toxic hazards and the possible benefits from risk reduction, social consequences and political values must play an integral role in determining which speculative risk estimates are adopted.

There is an inherent tension between the disciplinary norms of good science and good regulation. Unlike in pure scientific research, where the proper response to uncertainty is reservation of judgment pending the development of adequate data and testable hypotheses, the risk-assessment process cannot be suspended without significant social consequences. A finding that a vital issue is currently indeterminate would be entirely consistent with the practice of good science, but "no decision" on a possible toxic hazard inescapably is a decision that promotes interests which benefit from the regulatory status quo.[10] Risk assessment is not driven by the pursuit of knowledge for its own sake, the explicit goal of science, but by the need to decide whether potentially severe health hazards should be allowed to continue or whether high control costs should be imposed with potentially severe economic consequences. Thus, scientists in regulatory proceedings are expected to produce "answers" in a timely manner even if their predictions are highly speculative. Any reluctance to relax the stan-

ter OSHA Generic Cancer Policy]; Leape, *Quantitative Risk Assessment in Regulation of Environmental Carcinogens*, 4 HARV. ENVTL. L. REV. 86, 103 (1980).

7. *See* Cothern, Coniglio & Marcus, *Estimating Risk to Human Health*, 20 ENVTL. SCI. & TECH. 111, 113-15 (1986).

8. *Id.* at 115.

9. These speculative estimates are often presented in misleadingly precise quantitative terms. *See infra* text accompanying notes 91-144, 174-84.

10. *See* Bazelon, *Science and Uncertainty: A Jurist's View*, 5 HARV. ENVTL. L. REV. 209, 213 (1981); Latin, *supra* note 5, at 339.

dards of proof and certainty generally required of valid science may introduce a bias in favor of regulatory inaction.

Science aims at the dispassionate pursuit of truth. In contrast, scientists in risk-assessment proceedings frequently represent industries, labor unions, consumers, environmentalists, or agency bureaucracies with great interests at stake. These affiliations may often explicitly or unintentionally color interpretations of available evidence.[11] Scientists seldom base conclusions on data and experiments that cannot be reproduced, but information in regulatory hearings is routinely submitted by affected parties and frequently cannot be replicated or effectively challenged by other participants.[12] Scientists tend to design research studies in light of which data are available and which experiments may be feasible, whereas the critical questions in risk-assessment proceedings are usually determined by statutory or judicial requirements that need not be responsive to the state of scientific knowledge.[13] Budgetary and time limitations often influence the scientific research agenda, but no good scientist would feel that definitive answers must be produced irrespective of resource constraints. The opposite predisposition may be appropriate for good regulators.[14] These comments are not intended to call into question the competence or ethics of all scientists who participate in risk assessments. Rather, the point is that the risk-assessment process is fundamentally shaped by the requirements, constraints, and adversarial climate of regulation, not by the disciplinary norms of science.

The illusion that risk assessment is a purely scientific activity reduces the visibility and political accountability of policy judgments that often

11. For discussions of how conflicting private and bureaucratic incentives may impede effective environmental regulation, see Latin, *Ideal Versus Real Regulatory Efficiency: Implementation of Uniform Standards and "Fine-Tuning" Regulatory Reforms*, 37 STAN. L. REV. 1267, 1282–97 (1985); Stewart, *Regulation, Innovation, and Administrative Law: A Conceptual Framework*, 69 CALIF. L. REV. 1256, 1274–75, 1338–53 (1981). Scientists are no more immune to cognitive dissonance and wishful thinking than are nonscientists.

12. For example, agencies are largely dependent on polluting industries for information on current discharge levels and on the cost/profitability criteria needed to assess whether proposed standards would be economically feasible. *See* EPA Benzene Emissions from Maleic Anhydride Plants, Ethylbenzene/Styrene Plants, and Benzene Storage Vessels; Proposed Withdrawal of Proposed Standards, 49 Fed. Reg. 8386, 8389 (1984); Latin, *The Feasibility of Occupational Health Standards: An Essay on Legal Decisionmaking Under Uncertainty*, 78 NW. U.L. REV. 583, 605–11 (1983).

13. For examples of cases in which reviewing courts required quantitative risk assessments based on their interpretation of statutory provisions, without regard to whether the agencies were able to produce reliable risk estimates given the level of available data and scientific knowledge, see Industrial Union Dep't, AFL-CIO v. American Petroleum Inst., 448 U.S. 607 (1980); Gulf S. Insulation v. Consumer Prod. Safety Comm'n, 701 F.2d 1137 (5th Cir. 1983); Texas Indep. Ginners Ass'n v. Marshall, 630 F.2d 398 (5th Cir. 1980).

14. When harm will be substantially irreversible, as in the cases of carcinogenic exposures, extinction of species, or acid-rain contamination of lakes and forests, the problem of how long regulators should wait for "enough" information to enable reliable scientific judgments is likely to be controversial. *See* Latin, *supra* note 11, at 1282–83 & n.78; Latin, *supra* note 5, at 384–85.

guide regulatory decisions on toxic hazards. A comparison of conflicting risk-assessment principles adopted by agencies under different administrations shows that regulators frequently do consider policy criteria when they select specific risk estimates.[15] Federal agencies have recently employed controversial risk-assessment assumptions to justify inaction on some hazardous substances. Regulators have also attempted to make determinations based on "good science" without considering the implications of this approach for decisionmaking costs, regulatory delays, and opportunities for obstructive or strategic behavior by affected parties. Risk assessors often respond to scientific uncertainties by adopting conservative safety-oriented positions on some important issues while they use best-current-scientific-guess, middle-of-the-range, methodological-convenience, or least-cost treatments on other material issues. EPA and other agencies have never explained the scientific or policy rationales underlying these inconsistent treatments of uncertainty, and risk managers may not recognize that substantial inconsistency exists. In light of these diverse risk-assessment practices, I contend that regulatory policy judgments as well as scientific judgments must be applied coherently, explained forthrightly, and tested actively through public debate.

Several disparate reasons may explain the current emphasis on attaining "good science" in regulatory proceedings on toxic substances. After unsuccessful attempts to achieve environmental deregulation, the Reagan Administration adopted a strategy purportedly designed to improve the efficiency of pollution control programs.[16] One EPA Assistant Administrator contended that efficient standards must be based on "scientific evidence and not on rumor and soothsaying,"[17] and another official noted that the new cancer guidelines "hopefully will add to the scientific credibility" of agency decisions.[18] In a recent *Science* symposium on risk assessment, two EPA regulators claimed that the guidelines were intended "to reduce possible confusion by dealing consistently and openly with the assumptions and extrapolations that are required to bridge the gap between scientific findings and the risk assessments derived from them."[19]

Critics of the new approach regard Administration prescriptions for "good science" as a subterfuge designed to accomplish de facto deregulation. Dr. J. Donald Millar, the Director of the National Institute of

15. *See infra* text accompanying notes 44-52, 62-69, 198-202.

16. *See* Latin, *supra* note 11, at 1271-72.

17. Eidsness, *An Administration Sold on Clean Water*, N.Y. Times, Nov. 9, 1982, at 30, col. 4.

18. Shabecoff, *Administration Drafting New Policy on Regulating Cancer-Causing Agents*, N.Y. Times, Dec. 4, 1982, at 32, col. 2 (quoting Dr. Denis Prager, Assistant Director of the Office of Science and Technology).

19. Russell & Gruber, *Risk Assessment in Environmental Policy Making*, 236 SCIENCE 286, 286-87 (1987).

Occupational Safety and Health (NIOSH), observed that the previous carcinogen policy was being revised because "the chemical industry finds it to be an onerous responsibility."[20] Congressman, now Senator, Albert Gore, Jr., similarly argued: "The upper echelon science policy-makers have made a crass, calculated, cynical change in the traditional policy of seeking to prevent cancer."[21] He claimed that the Administration has "reached way down into the processes of government to control the science. They think that if you control the science you can control the conclusions about whether to control this or that substance."[22]

Both sets of characterizations may have some validity. Many scientists and regulators sincerely believe that the quality of risk assessments must be improved through application of the latest scientific findings, while requirements for "good science" may also be motivated by the Administration's outcome-oriented recognition that dispositive scientific evidence is unavailable on many contested issues. The critical fact is that this "good science" orientation, whatever its initial purposes, is becoming entrenched in a myriad of regulatory programs as agencies increasingly rely on quantitative risk assessment, risk-benefit analysis, or cost-benefit analysis[23] to justify pollution control decisions and to establish staff priorities.[24] EPA's carcinogen guidelines, for example, are likely to be the most influential statement of federal risk-assessment practices for years to come, and yet they have not been scrutinized from public policy and legal perspectives. It is important to stress that thousands of lives and billions of dollars in regulatory costs may depend on an agency's choice of controversial risk-assessment principles. The primary purpose of this Article is to encourage agency officials, legislators, and other legal decisionmakers to examine critically the scientific limitations and broader public policy implications of alternative risk-assessment treatments.

I. Risk-Assessment Principles in EPA's Carcinogen Guidelines

Social policy judgments have always been perceived as central to the risk-management process, and regulatory agencies have assigned different

20. Shabecoff, *supra* note 18, at 32, col. 2 (quoting J. Donald Millar).
21. Marshall, *EPA's High-Risk Carcinogen Policy*, 218 SCIENCE 975, 975 (1982).
22. *Id.*
23. A panel of the Court of Appeals for the District of Columbia Circuit, for example, recently approved EPA's controversial decision to employ cost-benefit analysis for hazardous air pollutant standards promulgated under §112 of the Clean Air Act. *See* Natural Resources Defense Council, Inc. v. EPA, 25 E.R.C. 1105 (D.C. Cir. 1986) (No. 85-1150). The Court of Appeals sitting *en banc* vacated the panel's opinion and heard oral arguments on April 29, 1987. On July 28, 1987, the *en banc* court remanded the case. 824 F.2d 1146 (D.C. Cir. 1987).
24. *See, e.g.*, Russell & Gruber, *supra* note 19, at 287-89; Lave, *Health and Safety Risk Analyses: Information for Better Decisions*, 236 SCIENCE 291 (1987).

weights to competing factors in response to changing political or economic conditions. Under the Carter Administration, risks above one fatality per million exposed people were usually treated as "unacceptable" if feasible control measures were available.[25] Reagan Administration agencies have concluded that risks as high as one in ten thousand, or even one in a hundred in some settings, are tolerable.[26] These risk-management decisions reflect different ideological preferences and different assumptions about the economic and political effects of toxic substances regulation. Similar considerations implicitly influence risk-assessment practices and resulting estimates of toxic hazards. Indeed, any decision by regulators to stress "good science" with the possible consequence of reduced public safety is itself a debatable policy choice. Yet, social policies and values adopted in risk-assessment proceedings typically have not been made explicit nor applied in a consistent manner.[27]

During the Carter Administration, EPA, OSHA, FDA, and the Consumer Product Safety Commission (CPSC) formed an Interagency Regulatory Liaison Group (IRLG) to develop a common set of risk-assessment principles. The IRLG guidelines, which were intended to achieve consistent resolutions of recurring scientific issues, emphasized the need for safety-oriented protective treatments under conditions of uncertainty.[28] OSHA and EPA also created generic cancer policies partly motivated by their desire to prevent repetitive submissions of scientific theories and supporting data that the IRLG had rejected as unreliable.[29] For example,

25. *See* Cross, *Beyond Benzene: Establishing Principles for a Significance Threshold on Regulatable Risks of Cancer*, 35 EMORY L.J. 1, 17 (1986). Carter Administration regulatory agencies often relied on qualitative analyses to demonstrate that a substance was toxic, but did not attempt to produce quantitative estimates of the risks associated with specific exposure levels.

26. *See id.* at 19-20; *The Odds on Cancer: EPA's Recent Bets*, 218 SCIENCE 976 (1982). For a specific example, see EPA Standards for Radon-222 Emissions from Licensed Uranium Mill Tailings, 51 Fed. Reg. 34,056, 34,057 (1986).

27. One caveat is necessary about the illustrations presented below. Each regulatory proceeding on toxic substances must address many complex scientific issues and requires analysis of great quantities of evidence. I believe useful evaluations of regulatory strategies cannot be performed without careful attention to specific factual circumstances and decisionmaking constraints. In other words, details matter and may often prove decisive. I have, for example, criticized some of the leading academics in the field of environmental law for advocating idealized theoretical approaches that underemphasize scientific uncertainties and practical implementation constraints on environmental regulation. *See* Latin, *supra* note 11, at 1273-1304, 1329-32. Yet, an extended discussion of risk-assessment problems and scientific data may submerge thematic points in a welter of technicalities. I have tried in this Article to attain a middle ground that neither oversimplifies difficult issues nor exhausts most readers' endurance, but some people will likely feel I erred in one direction or the other.

28. *See* Interagency Regulatory Liaison Group, Scientific Bases for Identification of Potential Carcinogens and Estimation of Risks, 44 Fed. Reg. 39,858 (1979) [hereinafter IRLG Guidelines].

29. *See* OSHA Generic Cancer Policy, 45 Fed. Reg. 5002 (1980); EPA National Emission Standards for Hazardous Air Pollutants; Policy and Procedures for Identifying, Assessing, and Regulating Airborne Substances Posing a Risk of Cancer, 44 Fed. Reg. 58,642 (1979). CPSC applied similar risk-assessment principles in its decision to regulate urea-formaldehyde foam insulation, see CPSC Ban of Urea-Formaldehyde Foam Insulation, 47 Fed. Reg. 14,366 (1982). The EPA generic policy

OSHA noted that industry representatives in every proceeding on toxic substances had argued for the existence of a threshold exposure level below which cancer risks are negligible.[30] The IRLG guidelines and OSHA generic cancer policy found this contention was unproven and implausible in biological terms. The agencies instead adopted a protective nonthreshold causation theory in recognition of continuing scientific uncertainty.[31] In the absence of any fundamental advance in the state of scientific understanding, OSHA and other IRLG agencies concluded that there was little reason to debate the threshold-level issue for every potential carcinogen.

Despite the Reagan Administration's generally dismal record on toxic substance regulation,[32] EPA's carcinogen guidelines may be examined at face value as an attempt to improve the quality and consistency of risk assessments. The specified practices usually conform to recommendations made by politically independent scientific organizations; the guidelines were widely reviewed by outside scientists; and in some instances the guidelines adopt conservative treatments similar to those in the IRLG guidelines.[33] There has, however, been a subtle but important shift in emphasis. Although the current guidelines are intended to encourage some degree of analytical consistency, EPA experts must now assess risks independently on the "weight of evidence" for each substance under review. The guidelines make clear that:

> [R]isk assessments will be conducted on a case-by-case basis, giving full consideration to all relevant scientific information. This case-by-case approach means that Agency experts review the scientific information on each agent and use the most scientifically appropriate interpretation to assess risk.[34]

cited here was proposed under the Carter Administration but was not adopted in final form by the Reagan Administration. The OSHA generic policy was formally adopted in 1980, but was subsequently ignored by agency staff under the new administration.

30. *See* OSHA Occupational Safety and Health Standards, 43 Fed. Reg. 5918, 5929 (1978).

31. *See id.* at 5946-47; OSHA Generic Cancer Policy, 45 Fed. Reg. 5002, 5023-24, 5131 (1980); IRLG Guidelines, 44 Fed. Reg. 39,858, 39,872-75 (1979).

32. *See, e.g.*, Cross, *supra* note 25; Latin, *supra* note 11, at 1309, 1324-29; Ashford, Ryan & Caldart, *A Hard Look at Federal Regulation of Formaldehyde: A Departure from Reasoned Decisionmaking*, 7 HARV. ENVTL. L. Rev. 297, 298-99, 330-31 (1983); Olson, *The Quiet Shift of Power: Office of Management & Budget Supervision of Environmental Protection Agency Rulemaking Under Executive Order 12,291*, 4 VA. J. NAT. RESOURCES L. 1 (1984).

33. *See* INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, 29 IARC MONOGRAPHS ON THE EVALUATION OF THE CARCINOGENIC RISK OF CHEMICALS TO HUMANS, SUPP. 4 (1982); NATIONAL RESEARCH COUNCIL, *supra* note 1; NATIONAL TOXICOLOGY PROGRAM, REPORT OF THE AD HOC PANEL ON CHEMICAL CARCINOGENESIS TESTING AND EVALUATION OF THE NATIONAL TOXICOLOGY PROGRAM (1984); Office of Science and Technology Policy, Chemical Carcinogens: Review of the Science and Its Associated Principles, 50 Fed. Reg. 10,372 (1985).

34. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992 (1986); *see* Albert, *supra* note 1, at 312-13.

The "weight of evidence" approach requires "an overall, balanced judgment of the totality of the available evidence"[35] that "should be dealt with on an individual basis."[36] This regulatory philosophy implies that risk assessors must examine any potentially relevant scientific theories and data that any party may choose to submit.[37] The guidelines never consider additional decisionmaking and administrative costs, regulatory delays, and opportunities for obstructive private behavior that may arise from implementation of this individualized "weight of evidence" treatment. It is fair to say that, in comparison with the IRLG approach, EPA now places considerably more stress on attempts to ground regulatory decisions on "good science" than on the need to provide effective pollution control under conditions of scientific uncertainty. Consider the following examples:

A. *Competing Extrapolative Models*

EPA's carcinogen guidelines follow the widely held view that "risks at low exposure levels cannot be measured directly either by animal experiments or by epidemiologic studies."[38] Analysts must therefore extrapolate from observed effects at high dosages to predicted risks at low exposure levels. They also must frequently extrapolate from results in high-dosage animal tests to animal risks and long-term human hazards at significantly lower doses. Scientists have developed a number of competing extrapolative models during the past two decades, but none has yet achieved general acceptance.[39] Although all of the models fit the observed high-dosage data reasonably well, their estimates of low-dosage hazards can vary by several orders of magnitude.[40] EPA's guidelines candidly acknowledge: "Goodness-of-fit to the [high-dose] experimental observations is not an effective means of discriminating among models."[41] In other words, there is usually

35. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,996 (1986).

36. *Id.* at 33,995; *see also id.* at 33,996-98, 34,001.

37. This individualized consideration specifically includes scientific claims on the threshold-level issue that would have been rejected by IRLG and OSHA. *See infra* text accompanying notes 43-52, 199-202.

38. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986); *see also* Albert, *supra* note 1, at 316.

39. For general discussions of competing extrapolative models, see OSHA Generic Cancer Policy, 45 Fed. Reg. 5002, 5184-88 (1980); CALIFORNIA DEP'T OF HEALTH SERVS., REPORT TO THE SCIENTIFIC REVIEW PANEL ON BENZENE: PART B—HEALTH EFFECTS OF BENZENE 68-80 (Nov. 1984) [hereinafter DHS BENZENE REPORT]; ENVIRON CORP., ELEMENTS OF TOXICOLOGY AND CHEMICAL RISK ASSESSMENT: A HANDBOOK FOR NONSCIENTISTS, ATTORNEYS AND DECISION MAKERS 37-41 (1986).

40. *See* EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986); Latin, *supra* note 5, at 371; Luken & Miller, *The Benefits and Costs of Regulating Benzene*, 31 J. AIR POLLUTION CONTROL ASS'N 1254, 1256-57 (1981).

41. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,998 (1986).

no reliable experimental basis for selecting one extrapolative model over another.[42] Dr. Peter Preuss, the present director of EPA's Office of Health and Environmental Assessment, observed in 1985:

> Until recently, risk assessments by Federal agencies generally utilized simple linear models which predicted that cancer risks were in direct proportion to the dose of carcinogen received; this uncomplicated methodology underscored the imprecise nature of risk calculations. Today, Federal agencies generally use complicated model-fitting computer programs in most calculations. The statistical techniques in these calculations are sophisticated and can be so involved that risk assessors themselves do not fully understand the details. The use of computerized models has led to some improvements in risk predictions, particularly in situations where the linear model does not fit experimental tumor data well. *However, the fundamental uncertainties arising from extrapolations between experimental animals and humans and from high doses to low have not significantly decreased.*[43]

The IRLG agencies adopted a "one-hit" linear extrapolative theory that assumed the absence of safe threshold levels; they did not, however, choose this approach simply because it was an "uncomplicated methodology." The linear one-hit model is the most conservative credible theory in the sense that it generates the highest risk estimates at low exposure levels. In an explicit policy judgment made in response to persistent uncertainties, the Carter Administration agencies chose to maximize safety at the possible cost of overly stringent regulation by adopting the most protective extrapolative model with significant support in the scientific community.[44]

The EPA guidelines recommend adoption of a linearized multistage model in most carcinogenic risk assessments.[45] This extrapolative theory is quite conservative and produces risk estimates at low exposure levels similar, though not usually equal, to the results of the one-hit linear model.[46]

42. *See id.* at 34,003; ENVIRON CORP., *supra* note 39, at 39.

43. Preuss & White, *The Changing Role of Risk Assessment in Federal Regulation*, 19 BANBURY REPORT, RISK QUANTITATION AND REGULATORY POLICY, 331, 335 (D. Hoel, R. Merrill & F. Perera eds. 1985) (emphasis added). At the time of this statement, Dr. Preuss was a senior scientist with the Consumer Product Safety Commission. EPA's Office of Health and Environmental Assessment is the division with primary responsibility for the Agency's choice of risk-assessment principles and for the conduct of most assessments.

44. *See, e.g.*, OSHA Generic Cancer Policy, 45 Fed. Reg. 5002, 5023-24 (1980); OSHA Occupational Exposure to Benzene, 43 Fed. Reg. 5918, 5946-47 (1978).

45. *See* EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997-98 (1986).

46. *See* ENVIRON CORP., *supra* note 39, at 40; Comments in 19 BANBURY REPORT, RISK QUANTITATION AND REGULATORY POLICY, 324, 327-28 (D. Hoel, R. Merrill & F. Perera eds. 1985).

EPA selected the multistage model on the grounds that it provides a better fit with the available experimental evidence than the one-hit model and also appears more compatible with current knowledge about some biological processes related to cancer causation.[47] Thus, the Agency adopted a protective, but not worst-case, extrapolative theory because it considers the multistage model most plausible based on the present state of scientific understanding. It is not, however, apparent why the Agency should prefer marginally greater scientific plausibility to marginally greater public protection given EPA's recognition that no extrapolative model is demonstrably correct and that goodness-of-fit for high-dose results does not prove a model's value in predicting low-dose effects.[48] The multistage theory may be tenable science in light of our imperfect knowledge about carcinogenesis mechanisms,[49] but EPA's selection of this provisional extrapolative model in pursuit of "good science" represents an implicit social policy judgment.

Moreover, the guidelines make clear that the linearized multistage model is a default methodology to be used "[i]n the absence of adequate information to the contrary."[50] Agency experts or regulated parties may now argue for adoption of competing models on the basis of individualized circumstances.[51] The guidelines provide no selection criteria for competing extrapolative theories in specific circumstances, and simply state: "When a different model is chosen, the risk assessment should clearly discuss the nature and weight of evidence that led to the choice."[52] This treatment gives broad, if not unlimited, discretion to Agency analysts and encourages regulated parties to present any extrapolative theories and data that support the outcome they desire.

B. *Aggregation of Benign and Malignant Tumors*

The current guidelines indicate: "Benign tumors should generally be combined with malignant tumors for risk estimates unless the benign tumors are not considered to have the potential to progress to the associ-

47. *See* DHS BENZENE REPORT, *supra* note 39, at 79-80; Albert, *supra* note 1, at 316.

48. *See supra* text accompanying notes 40-42.

49. *See* EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986). "At present, mechanisms of the carcinogenesis process are largely unknown and data are generally limited." *Id.*

50. *Id.*

51. *See id.* at 33,997-98, 34,003; Albert, *supra* note 1, at 316-17.

52. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,998 (1986). Individualized circumstances may be established by longitudinal dose-response data, pharmacokinetic and metabolism data, or "other substantial evidence on the mechanistic aspects of the carcinogenesis process." *Id.* The passage continues: "Considerable uncertainty will remain concerning response at low doses; therefore, in most cases an upper-limit risk estimate using the linearized multistage procedure should also be presented." *Id.* There is no indication in the guidelines of which circumstances would justify omission of an estimate derived from the multistage model.

ated malignancies of the same histogenic origin."[53] Unlike under the prior IRLG practice, toxic dischargers can now argue that the "weight of evidence" demonstrates particular benign tumors are not related to the onset of malignancy. When Dr. Roy Albert, the Chair of EPA's Carcinogen Guideline Committee and Carcinogen Assessment Group (CAG), was asked to describe the circumstances under which such an exception might be appropriate, he conceded: "I doubt if there is any persuasive rationale to exclude benign tumors. Benign tumors may represent the weak action of a carcinogen or be a signal that the agent is a promoter."[54] EPA contends that its treatment "allows flexibility in evaluating the data base for each agent,"[55] but the guidelines offer no criteria that would constrain exclusion of benign tumors.[56] In return for a possibly chimerical attempt to ensure the best scientific assessment in any circumstances, this individualized exception may increase the costs of the risk-assessment process and is likely to facilitate submissions intended to delay regulation.

C. *Negative Epidemiological Studies*

Industry typically submits epidemiological studies that purport to demonstrate no significant hazard at the low exposure levels currently prevalent in most toxic contexts. The IRLG agencies consistently rejected this type of evidence on the grounds that no level of exposure to a known carcinogen could be entirely safe, and that the negative studies invariably suffered from serious methodological deficiencies.[57] In its 1984 proceedings on benzene emissions from coke by-product recovery plants, EPA refused to accept negative epidemiological findings submitted by industry for the same reasons that other agencies had previously rejected them.[58] The Agency's current carcinogen guidelines, in contrast, provide:

> It should be recognized that epidemiologic studies are inherently capable of detecting only comparatively large increases in the relative risk of cancer. Negative results from such studies cannot prove the absence of carcinogenic action; however, negative results from a well-designed and well-conducted epidemiologic study that contains usable exposure data can serve to define upper limits of risk; these are

53. *Id.* at 33,997.
54. Comments, *supra* note 46, at 324.
55. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 34,002 (1986).
56. *See id.* Instead, the Agency stated generally that scientific issues should be addressed on their individual merits in light of the available evidence. *See supra* text accompanying notes 35-36.
57. See Gulf S. Insulation v. Consumer Prod. Safety Comm'n, 701 F.2d 1137, 1146 (5th Cir. 1983). *See* OSHA Identification, Classification and Regulation of Toxic Substances Posing a Potential Occupational Carcinogenic Risk, 42 Fed. Reg. 54,148, 54,155-56 (1977); OSHA Occupational Exposure to Benzene, 43 Fed. Reg. 5918, 5928-31, 5946 (1978);
58. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,479, 23,484 (1984).

useful if animal evidence indicates that the agent is potentially carcinogenic in humans.[59]

Reliable negative epidemiologic data on carcinogenic risks will seldom be available from "well-designed and well-conducted" studies. Many cancers entail latency periods of several decades, and the period may sometimes increase for smaller doses.[60] Thus, accurate data must be available from past years when exposures were frequently much higher than currently permissible and when exposure monitoring was typically casual at best. Moreover, negative findings cannot be meaningful without a large enough cohort to present a statistically valid sample, without accurate exposure data for cohort members, without a reliable follow-up of subjects who leave the place of employment or other cohort, and without accurate determinations of the actual causes of death. These methodological characteristics are extremely uncommon in long-term studies of human exposures to chemical substances.[61] Despite these barriers to implementation of a "well-designed and well-conducted epidemiologic study," the EPA guidelines allow industry to submit any negative studies in an attempt to "define upper limits of risk" and consequently require Agency scientists in each regulatory proceeding to examine those problematical studies in detail.

D. *Body-Site and Tumor-Type Specificity*

The guidelines conservatively provide that "the biologically acceptable data set from long-term animal studies showing the greatest sensitivity should generally be given the greatest emphasis."[62] However, they also maintain that a "statistically significant excess of tumors of all types in the aggregate, in the absence of a statistically significant increase of any individual tumor type, should be regarded as minimal evidence of carcinogenic action unless there are persuasive reasons to the contrary."[63] To determine the overall risk, officials must add together all tumor types or sites that appear at a significant level, but "quantitative risk extrapolations will

59. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,995-96 (1986).

60. *See* OSHA Generic Cancer Policy, 45 Fed. Reg. 5002, 5040-41 (1980).

61. *See, e.g.*, EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,479 (1984); Latin, *supra* note 5, at 361-65, 370 & n.257.

62. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986). The Agency selected this protective treatment "[b]ecause it is possible that human sensitivity is as high as the most sensitive responding animal species." *Id.* In contrast, some scientists contend that results from all relevant animal studies should be averaged. *See* Comments, *supra* note 46, at 325-26. As usual, the EPA guidelines direct that data from the most sensitive species should be used "in the absence of evidence to the contrary," without specifying what that individualized evidence might be. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986).

63. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,995 (1986).

generally not be done on the basis of totals that include tumor sites without statistically significant elevations."[64] In other words, contrary to the prior IRLG treatment,[65] EPA ordinarily will require proof of a statistically significant number of specific tumors in specific organs or body locations. A statistically significant total of tumors of varying types in varying locations normally will not be deemed a sufficient basis for regulation.

There is a tenable scientific rationale for this treatment[66] but it is not necessarily mandated by good science. At a recent conference on risk assessment, one scientist critical of EPA's current practice noted that some carcinogenic agents are "just not very organ-specific."[67] Dr. Albert defended the Agency treatment as "an attempt at a middle-of-the-road position."[68] EPA has never explained why a "middle" position is desirable on a safety issue where the Agency acknowledges the absence of a clear scientific consensus.[69] EPA's selection of a mid-range position on this issue reflects an implicit social policy choice that is not required by the norms of good science and that cannot be resolved solely on the basis of scientific judgments.

E. *Extrapolation from Animal Test Dosages*

Reliable epidemiologic data on long-term human risks cannot exist for recently introduced toxic substances and is rarely available for chemicals in longstanding use. Thus, most attempts to regulate toxic hazards must rely on extrapolation of human risks from animal test data. Because rodents and most other experimental subjects have much smaller bodies than humans, risk assessors must adopt an "interspecies scaling factor" to link test dosages with corresponding human exposures. EPA's carcinogen guidelines only briefly explain its treatment of this issue:

> The usual approach for making interspecies comparisons has been to use standardized scaling factors. Commonly employed standardized dosage scales include mg per kg body weight per day, ppm in the diet or water, mg per m^2 body surface area per day, and mg per kg

64. *Id.* at 33,997; *see also id.* at 34,003.

65. *See* Ashford, Ryan & Caldart, *supra* note 32, at 298–99, 330–31; Latin, *supra* note 11, at 1325, 1328.

66. *See* the observations of Drs. Albert, Purchase, and Weinstein in Comments, *supra* note 46, at 324–25. Inclusion of all observed tumors may lead to overestimation of risks if some result from background cancer rates or other causes unrelated to the toxic exposures under review. On the other hand, toxic risks may be underestimated if some cancers detected in animal tests or epidemiologic data are excluded despite a statistically significant incidence of total tumors.

67. *Id.* at 324 (comments of Dr. Weinstein); *see also id.* at 324–25.

68. *Id.* at 325.

69. *See id.* (comments of Dr. Albert) ("It is perfectly clear that here, in this meeting, we have views that range from describing this sort of evidence as 'poor' to others as 'not so bad'.").

> body weight per lifetime. In the absence of comparative toxicological, physiological, metabolic, and pharmacokinetic data for a given suspect carcinogen, the Agency takes the position that the extrapolation on the basis of surface area is considered to be appropriate because certain pharmacological effects commonly scale according to surface area.[70]

This provision indicates that Agency risk assessors must consider the individualized scientific evidence associated with each toxic substance before selecting an interspecies scaling factor. The guidelines never examine the impacts of this treatment on Agency decisionmaking costs and on the ability of affected parties to challenge the scientific bases for risk estimates derived from animal studies.

Moreover, EPA chose its "default" scaling factor on the ground that "certain pharmacological effects commonly scale according to surface area," but the Agency made no attempt to show that those effects bear a reasonable relationship to any carcinogenesis process or that a scientific consensus supports this approach. The significance of this default treatment may be put in perspective by considering alternative estimates presented in a California Department of Health Services (DHS) study of animal test data on benzene-related risks.[71]

TABLE 1

Interspecies Scaling Factor	Human Risk Per PPB
mg/kg of body wt. per day	14×10^{-6}
mg/body surface area per day	170×10^{-6}
mg/kg of body wt. per lifetime	580×10^{-6}

DHS offered these benzene risk estimates only "for illustrative purposes" because "no study has been explicitly undertaken with the objective of determining what unit best expresses equivalence of carcinogenic potency across mammalian species."[72] The Agency then followed EPA's choice of the "middle of the range" surface-area criterion.[73] Absent any discussion of the relationship between "certain pharmacological effects" and carcinogenesis processes, it is unclear whether EPA selected the body-surface-area criterion because it represents the best available current science or because it leads to a mid-range assessment of toxic

70. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,998 (1986) (citations omitted).
71. DHS used data from studies of male mouse preputial gland cancers to calculate the human benzene risks presented in the accompanying table. DHS BENZENE REPORT, *supra* note 39, at 106.
72. *Id.* at 68.
73. *Id.*

risks. The choice of scaling factor may produce risk estimates varying by more than an order of magnitude and yet EPA has never explained why its mid-range default position is appropriate under conditions of scientific uncertainty.

The discussions above are not intended to suggest that regulatory agencies should always provide the greatest conceivable degree of protection against toxic risks under conditions of uncertainty. Rather, they demonstrate that EPA's carcinogen guidelines contain numerous treatments that are not compelled by generally accepted norms of risk-assessment science and that have policy implications which cannot be evaluated purely in scientific terms. Whether the likely ramifications from EPA's pursuit of "good science" constitute a serious problem depends on how reliable current risk-assessment practices really are, for few people would object to basing regulatory judgments on *accurate* analyses of toxic hazards. A more detailed description of the scientific uncertainties that arose in a specific risk-assessment proceeding may be useful to show how tenuous, one might even say flimsy, is the technical evidence on which major pollution control actions (or inaction) are often grounded.

II. EPA Analyses of Benzene Risks

This Part of the Article presents an extended description of scientific issues raised in EPA's 1979 and 1984 regulatory proceedings on benzene, one of the most common toxic substances found in the workplace and ambient environment. Indeed, it is probably the most widely used industrial chemical for which a carcinogenic causal relationship has been clearly demonstrated. Benzene has consequently been the subject of numerous proceedings by EPA, OSHA, CPSC, and state air pollution control authorities. The following discussion of EPA's regulatory decisions does not attempt to examine all, or even a majority, of the technical issues and uncertainties presented in benzene proceedings, but it nonetheless cites many complex and often tedious details that regulators inevitably must confront. Before turning to specific scientific issues, it is necessary to explain briefly why familiarity with the details of actual risk-assessment practices is essential for the development of effective legal and social policies in this field.

A. *The Need for Particularized Implementation Analyses*

Most treatments of risk-assessment issues in the legal literature are highly conceptual and are dominated by *regulatory paradigms* rather than insights drawn from *regulatory practice.* These legal or policy-

oriented discussions generally contend that perfect safety is unattainable,[74] that agency decisionmakers should balance numerous incompatible and often incommensurable criteria to reach the most efficient possible outcome,[75] and that it is essential for regulators to "ask the right questions" despite the presence of scientific uncertainty.[76] In contrast to these abstract generalizations, I believe useful analyses of risk-assessment practices require a realistic consideration of implementation constraints, administrative costs, and bureaucratic incentives associated with regulation in a field where dispositive answers are seldom available. I have previously noted, for example, that a typical consequence of requiring regulators to address currently unanswerable scientific questions is agency paralysis, not improved decisionmaking.[77]

In response to my emphasis on implementation problems, Bruce Ackerman and Richard Stewart contended that environmental regulation must be made more efficient by adopting a "pollution-based" cost-effectiveness approach that "would encourage a more focused discussion of whether the goals set for different pollutants reflect sensible priorities."[78] They then continued:

> Indeed, it is not *fanciful* to suppose that a risk portfolio strategy eventually might emerge that would explicitly attempt to rank the comparative risks confronted by an EPA or an OSHA and then use cost-effectiveness analysis to determine how available administrative and control resources might best be devoted to minimizing overall risk in a given time period. Such a strategy need not be limited to conventional air and water pollutants. It could also be used, for example, to manage the risks posed by pesticides, chemicals, or hazardous wastes.[79]

The footnote to this passage further claimed that, to reduce administrative burdens, "a 'mutual fund' variant of the portfolio approach might be used, where appropriate, to control related pollutants through permits

74. *See, e.g.*, M. DOUGLAS & A. WILDAVSKY, RISK AND CULTURE (1982); Huber, *Safety and the Second Best: The Hazards of Public Risk Management in the Courts*, 85 COLUM. L. REV. 277 (1985); Pedersen, *Why the Clean Air Act Works Badly*, 129 U. PA. L. REV. 1059 (1981).

75. *See, e.g.*, B. ACKERMAN & W. HASSLER, CLEAN COAL/DIRTY AIR, 79–103 (1981); D. CURRIE, AIR POLLUTION: FEDERAL LAW AND ANALYSIS §§ 7.13, 10.01 (1981); Krier, *The Irrational National Air Quality Standards: Macro- and Micro-Mistakes*, 22 UCLA L. REV. 323, 324–30 (1974).

76. *See, e.g.*, B. ACKERMAN & W. HASSLER, *supra* note 75, at 103; Ackerman & Stewart, *Reforming Environmental Law*, 37 STAN. L. REV. 1333, 1357 (1985).

77. *See* Latin, *supra* note 11, at 1281–84, 1329–30.

78. Ackerman & Stewart, *supra* note 76, at 1360.

79. *Id.* at 1360–61 (footnotes omitted) (emphasis added).

based on a weighted average of volume and risk."[80] Implementation of these proposals, which I regard as not merely "fanciful" but fantastic, would clearly require a sophisticated capability for reliable risk assessments. Regulators could hardly trade the risks posed by one toxic substance against the risks from other toxics in the "portfolio" unless they can ascertain the relative hazards with reasonable confidence. Professors Ackerman and Stewart minimized the significance of scientific uncertainty by claiming that I "underestimate the amount of information which *does* exist, but which is ignored by regulators who refuse to confront ecological and economic realities."[81] As support for this claim, they cited recent decisions by EPA not to regulate "acrylonitrile and other toxic emissions . . . when analysis suggested that the risks involved were relatively low compared to the more serious problems posed by chromium emissions—which it did decide to regulate."[82] Ackerman and Stewart contended these "examples show that it is feasible to do a better job of goal-setting—by introducing cost-effectiveness considerations in evaluating control options for different risks and setting priorities more intelligently."[83]

This "proof" is unpersuasive because it assumes as a given that EPA's risk assessments were accurate and that the magnitude of the "less serious" risks did not warrant regulation. The bare fact that EPA decided to regulate one substance but not others cannot support a conclusion that the Agency's decisions were reliable from the perspective of science or efficient from the perspective of social policy. Certainly, agencies must set regulatory priorities in some manner, but this truism does not suggest that current risk assessment practices fairly reflect prevailing social and political values. I believe Ackerman and Stewart's contention that EPA did a "better job of goal-setting" and acted "more intelligently" in the regulatory contexts they cited was not grounded on a careful analysis of specific risk assessments, but was only wishful thinking used to justify their idealized conceptualizations. The critical risk-assessment issues—whether current treatments yield reasonably accurate risk estimates and whether agency practices implicitly incorporate controversial social policy judgments concealed beneath a veneer of "good science"—cannot be resolved in the abstract. Rather, a careful examination of actual risk-assessment practices is essential for the selection of effective regulatory approaches.

EPA proceedings on benzene examined much more scientific and medical evidence than was available for acrylonitrile and the other chemicals

80. *Id.* at 1360 n.62.
81. *Id.* at 1363 (emphasis in original).
82. *Id.*
83. *Id.*

cited by Ackerman and Stewart. Indeed, I chose to review EPA's treatment of benzene risks largely because of the relative "abundance" of scientific information and the relative "depth" of the Agency's considerations.[84] Moreover, the analytical practices adopted in the benzene proceedings are very similar to the risk-assessment principles later incorporated in EPA's carcinogen guidelines. Although each regulatory analysis of toxic hazards must be evaluated on its individual merits, there is little doubt that most of the problems described below were equally or more severe in the EPA risk assessments uncritically accepted by Ackerman and Stewart.

I have tried to keep as succinct as possible my discussions of scientific issues raised in EPA's benzene proceedings; yet, considerable detail is essential because descriptions of actual regulatory problems are the only antidote to the blithe conceptualizations that pervade most legal analyses of risk-assessment issues. Because some readers will find this level of particularity burdensome, while others will appreciate specific details, it may be useful to list in advance the principal observations documented below:

1. Different regulatory agencies adopted different risk-assessment principles, relied on different types of scientific evidence, and reached different conclusions about the extent of benzene-related health hazards.
2. Despite widespread use of benzene for decades and ample evidence that the substance is carcinogenic, there was little available data on the hazards associated with specific exposure levels.
3. Risk assessors adopted conservative positions on some issues but employed best-current-guess, mid-range, or methodological-convenience treatments on other important issues. EPA presented no explanations for these disparate treatments of scientific uncertainty.
4. On many risk-assessment issues, EPA acknowledged that no scientific consensus exists and that the principal uncertainties cannot be resolved in anything approaching a reliable manner.
5. EPA frequently dismissed material uncertainties by pairing discussions of indeterminate factors that could lead to overestimation and underestimation of possible hazards, with no attempt to compare the relative significance of those ostensibly offsetting factors.
6. EPA and other agencies often presented risk estimates with an implausible degree of apparent precision. They did not, however, agree on how estimated risks should be described: Some assessments presented the risks as discrete point estimates, while others cited ranges of possible risks at varying exposure levels.

84. I also had the advantage of familiarity with many of the relevant technical issues because I previously evaluated OSHA's attempt to regulate benzene as an occupational health hazard. *See* Latin, *supra* note 5.

7. Industry representatives submitted the same debatable scientific interpretations and unvalidated data in each benzene proceeding, but EPA and other agencies seldom explicitly considered the decision-making costs and regulatory delays imposed by individualized considerations of repetitive arguments and evidence.

Readers who are prepared to accept these observations and who would find a detailed analysis of diverse scientific issues unproductive may prefer to read only subsection E in this Part and then the more general discussions of problems raised by EPA's current "good science" approach and of potentially applicable social policies in Parts III and IV.

B. *EPA's Reliance on Epidemiologic Data to Derive a Unit Risk Factor for Benzene*

In 1978 OSHA regulated benzene as an occupational carcinogen but decided there was insufficient evidence to identify specific benzene risks at varying exposure levels.[85] OSHA never produced a quantitative risk assessment because it found the epidemiologic data were not accompanied by accurate exposure measurements and because it questioned the reliability of existing scientific models for extrapolation from data on high-dose effects to predictions about low-dose risks.[86] In an explicit social policy judgment, OSHA concluded that it should not wait for scientific certainty[87] while hundreds of thousands of workers were exposed to benzene levels that could feasibly be reduced.[88]

EPA, in contrast, has twice produced quantitative estimates of the environmental risks from benzene discharges. Based on a 1979 analysis by its Carcinogen Assessment Group (CAG), EPA proposed to regulate emis-

85. In 1969 benzene dischargers adopted by industry consensus an occupational exposure limit of 10 parts per million (ppm) in response to evidence that benzene causes nonmalignant blood-disorders. OSHA accepted this 10 ppm permissible exposure limit (PEL) in a 1971 national consensus standard. During the next few years, OSHA accumulated epidemiologic evidence from several studies showing that benzene causes leukemia. Based on evidence of excess leukemia deaths resulting from high or unknown historical exposures, the Agency made a qualitative determination that benzene is an occupational carcinogen. The Agency then attempted in 1978 to reduce the PEL from 10 ppm to 1 ppm. *See* Industrial Union Dep't, AFL-CIO v. American Petroleum Inst., 448 U.S. 607, 617-28 (1980); OSHA Occupational Exposure to Benzene, 43 Fed. Reg. 5918, 5921-25 (1978); Latin, *supra* note 5, at 344-46. This regulatory standard was invalidated by the Supreme Court on the ground that OSHA had failed to prove prevailing exposure levels pose a "significant" risk of harm. *See Industrial Union*, 448 U.S. at 607. For a brief discussion of this judicial treatment, see *infra* text accompanying notes 210-22.

86. *See Industrial Union*, 448 U.S. at 635-36; Latin, *supra* note 5, at 359-71; *see also* OSHA Generic Cancer Policy, 45 Fed. Reg. 5002, 5196-200 (1980).

87. OSHA Occupational Exposure to Benzene, 43 Fed. Reg. 5918, 5920 (1978).

88. OSHA selected the 1 ppm PEL not because it was entirely safe, but because it was the lowest discharge level for which emissions controls were technologically and economically feasible. *See Industrial Union*, 448 U.S. at 621 & n.14, 650-51; Latin, *supra* note 5, at 644.

sions from several categories of industrial sources.[89] In 1984, EPA withdrew all but one proposed standard on the ground that affected industries had voluntarily reduced discharges to the point where residual risks were no longer "significant."[90] The Agency did promulgate a revised standard for emissions from coke by-product recovery plants after predicting the following risks:

TABLE 2

EPA's 1984 Estimate of Leukemia Risks From All Coke By-Product Recovery Plants[91]

Risks of Leukemia from Lifetime (70 years) Exposures To Benzene	Number of Exposed People At Risk Within 20 km of Sources
1 in 100 to 1 in 1,000	3,200
1 in 1,000 to 1 in 10,000	101,000
1 in 10,000 to 1 in 100,000	2,212,000
1 in 100,000 to 1 in 1,000,000	17,991,000
1 in 1,000,000 to 1 in 10,000,000	10,214,000
1 in 10,000,000 to 1 in 100,000,000	442,000

EPA derived these risk figures by multiplying the estimated lifetime probability of leukemia for each part per million (ppm) of exposure by the number of people exposed to different atmospheric benzene levels. This mode of risk assessment entails three primary variables: the "unit risk factor" or lifetime chance of contracting leukemia per dosage unit, the ambient benzene concentrations surrounding each source, and the number of people in proximity to the sources.[92] Based on a predicted unit risk

89. *See* EPA National Emission Standard For Hazardous Air Pollutants; Benzene Emissions From Maleic Anhydride Plants, 45 Fed. Reg. 26,659 (1980); EPA National Emissions Standards For Hazardous Air Pollutants; Benzene Emissions From Ethylbenzene/Styrene Plants, 45 Fed. Reg. 83,447 (1980); EPA Benzene Emissions From Benzene Storage Vessels; National Emission Standard for Hazardous Air Pollutants, 45 Fed. Reg. 83,591 (1980); EPA National Emission Standard for Hazardous Air Pollutants; Benzene Fugitive Emissions, 46 Fed. Reg. 1165 (1981); Cross, *supra* note 25, at 26-27. EPA eventually issued the proposed standard for fugitive benzene emissions, but the Agency took that action only after a court order required it. *See* EPA National Emission Standards for Hazardous Air Pollutants; Benzene Equipment Leaks (Fugitive Emission Sources), 40 C.F.R §§61.110-.112 (1987).

90. *See* EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558 (1984); EPA Benzene Emissions from Maleic Anhydride Plants, Ethylbenzene/Styrene Plants, and Benzene Storage Vessels; Proposed Withdrawal of Proposed Standards, 49 Fed. Reg. 8386 (1984). The Consumer Product Safety Commission similarly withdrew its proposed benzene standard on the grounds that manufacturers had voluntarily chosen to eliminate most intentional uses of benzene in product designs. *See* CPSC Benzene-Containing Consumer Products; Proposed Withdrawal of Proposed Rule, 46 Fed. Reg. 3034 (1982). Thus, there is no applicable CPSC standard if manufacturers decide in the future to increase the benzene content of their products.

91. EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,527 (1984).

92. *See id.* at 23,525-26; EPA National Emission Standards for Hazardous Air Pollutants; Benzene Emissions from Maleic Anhydride Plants, Ethylbenzene/Styrene Plants, and Benzene Storage

factor of 0.022 leukemias per ppm of benzene[93] and on emissions monitoring data from some coke by-product recovery plants, EPA concluded that only 3200 people will be exposed to high concentrations subjecting them to a lifetime leukemia risk of greater than one in a thousand. EPA estimated that the annual incidence of benzene-induced leukemias for this "most exposed population" group is 2.2 cases.[94] Many millions of other people are exposed to lower concentrations that entail estimated lifetime risks of 10^{-4} to 10^{-8} leukemias.[95]

The 1979 CAG analysis depended almost entirely on linear extrapolations from epidemiologic data in three studies. Despite minor revisions of CAG findings in response to criticisms submitted by industry, EPA's final 1984 estimates were grounded on the same data and analytical assumptions used by CAG in 1979. A review of the epidemiologic evidence demonstrates that the data were insufficient to support reliable scientific conclusions and that the Agency adopted a number of debatable assumptions to bridge the large analytical gaps.

1. *The Aksoy Study*

Based on studies at two hospitals in Istanbul, Dr. Aksoy and his associates concluded that workers in unventilated shoe and handbag manufacturing shops were subject to a risk of 13×10^{-5} leukemias in comparison with a 6×10^{-5} background risk for the general population.[96] He identified 34 leukemia cases among 28,500 exposed workers, but this worker-population figure was derived from rudimentary "official records" rather than detailed industrial hygiene data.[97] Aksoy did not possess individualized data on actual exposure levels. He assumed after a few sample measurements that workers were exposed to maximum concentrations of 150 to 210 ppm and, because they often lived in their shops, to average con-

Vessels; Proposed Withdrawal of Proposed Standards, 49 Fed. Reg. 8386, 8387 (1984); Albert, *supra* note 1, at 317-19.

93. *See* EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,527 (1984).

94. *Id.*

95. *See id.* at 23,525-27; EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,489 (1984). I find it surprising and unfortunate that EPA never presented a cumulative estimate of predicted benzene-related fatalities for all exposed population groups.

96. *See* R. ALBERT, CARCINOGEN ASSESSMENT GROUP'S FINAL REPORT ON POPULATION RISK TO AMBIENT BENZENE EXPOSURES 13–14 (Jan. 10, 1979) [hereinafter CAG BENZENE REPORT]. Aksoy estimated the background lifetime leukemia risk as 6/100,000, but other experts submitted estimates ranging from 14×10^{5} down to 3×10^{5} cases. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984). The higher background risk figure would eliminate the statistical significance of Aksoy's findings about benzene-induced effects, while adoption of the lower background estimate would double the relative risk predicted by the CAG. *Id.*

97. CAG BENZENE REPORT, *supra* note 96, at 13; *see also* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984).

centrations of 15 to 30 ppm outside of working hours.[98] The CAG used these tentative exposure estimates to calculate a unit risk factor of .020252 per ppm.[99]

The critical risk-assessment issue is not whether the Aksoy study supports a qualitative finding that benzene exposure causes excess leukemia deaths—though industry disputed the study's utility even for that purpose—but whether Aksoy's data were sufficient to derive a reliable dose-response curve for low-level exposures. Virtually every quantitative estimate and analytical assumption in the study was challenged in the scientific literature or the regulatory proceedings.

For example, CAG used the geometric mean of Aksoy's high and low exposure estimates (between 150-210 and 15-30 ppm) to derive an average of 63.6 ppm per working hour.[100] Yet, the International Agency for Research on Cancer (IARC) reported that peak exposures had been reported as high as 650 ppm.[101] Moreover, there was no assurance that only 28,500 workers were exposed to benzene nor that all included workers were subjected to relatively high emissions levels.[102] Dr. Bernard Goldstein, EPA's Assistant Administrator for Research and Development, noted that Aksoy's "scanty exposure data" consisted only "of a few samples with no information as to how these samples may have typified the average benzene exposure in the workplace."[103]

The IARC evaluation also noted that the type of cancer clearly associated with benzene exposure, acute non-lymphocytic leukemia, appeared far more often in exposed workers than in the general population. IARC concluded that if the incidence in Aksoy's data of this specific type of leukemia, rather than all leukemias, had been used to compute the relative risks from benzene exposure, the excess risk would be more than an order of magnitude greater than the CAG prediction.[104]

In its 1984 explanation of the benzene standard, EPA acknowledged the validity of several methodological criticisms that would decrease the

98. *See* CAG BENZENE REPORT, *supra* note 96, at 16.

99. *See id.* at 17. CAG based this conclusion on a lifetime average exposure estimate of 4.22 ppm, which it derived using an average ten-hour working day and a 300-day working year. *See id.* at 16-17. This treatment may underestimate lifetime exposures, and hence overestimate the risk associated with lower emissions levels, because many of the workers were subject to continuing benzene exposures in their shared working-living quarters. *See* Luken & Miller, *supra* note 40, at 1256.

100. *See* CAG BENZENE REPORT, *supra* note 96, at 16.

101. *See* INTERNATIONAL AGENCY FOR RESEARCH ON CANCER, 29 IARC MONOGRAPHS ON THE EVALUATION OF THE CARCINOGENIC RISK OF CHEMICALS TO HUMANS 121 (1982) [hereinafter IARC BENZENE STUDY].

102. *See id.*; EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984).

103. *See* Goldstein, *Risk Assessment and Risk Management of Benzene by the Environmental Protection Agency*, 19 BANBURY REPORT, RISK QUANTITATION AND REGULATORY POLICY, 293, 295 (D. Hoel, R. Merrill & F. Perera eds. 1985).

104. *See* IARC BENZENE STUDY, *supra* note 101, at 122.

predicted levels of risk in Aksoy's findings. The Agency observed: "It is equally likely, however, that Aksoy's methodology leads to an underestimate of the excess risk."[105] EPA then summarized several plausible criticisms that would increase the estimate of possible risks, but the Agency made no attempt to assess the magnitude of these conflicting sets of criticisms. Instead, it apparently hoped that overestimated and underestimated risk factors would somehow counterbalance each other, a treatment hardly consistent with the practice of reliable science. Given the numerous uncertainties underlying the Aksoy data and the CAG analysis, EPA's presentation of the predicted .020252/ppm unit risk factor to a precision of six decimal places was inherently misleading.[106]

2. *The Ott Study*

Dr. Ott and his colleagues examined a cohort of 594 chemical manufacturing workers exposed to benzene between 1940 and 1973. They identified three cases of leukemia compared to an expected level of 0.8, which represents an excess risk of 3.75.[107] CAG used Ott's exposure and risk estimates to derive a unit risk factor of .04638/ppm of benzene.[108] EPA's reliance on the Ott study has been criticized on several grounds:

CAG characterized Ott's finding of three leukemias in 594 workers as of "borderline statistical significance."[109] EPA nonetheless relied on Ott's data in 1984 because it was the only study that involved relatively low levels of benzene exposure.[110]

CAG accepted Ott's attempt to estimate exposures for different job categories "as accurately as the historical air monitoring data permitted."[111] Yet, these estimates were based on work history data and plant hygiene measurement surveys, not on personalized records of benzene exposure.[112] IARC observed that "the number of workers in any particular work area was limited, and the power of the study to detect any association between exposure levels and cases was correspondingly low."[113]

Industry experts noted that one of the three identified leukemias had

105. *Id.*

106. *See infra* text accompanying notes 131–32, 162–84.

107. The excess risk was derived by dividing the observed incidence of 3 leukemias by the expected incidence of 0.8. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984); CAG BENZENE REPORT, *supra* note 96, at 18.

108. *See* CAG BENZENE REPORT, *supra* note 96, at 19–20.

109. *Id.* at 18.

110. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984); EPA Benzene Emissions from Maleic Anhydride Plants, Ethylbenzene/Styrene Plants, and Benzene Storage Vessels; Proposed Withdrawal of Proposed Standards, 49 Fed. Reg. 8386, 8388 (1984).

111. *See* CAG BENZENE REPORT, *supra* note 96, at 17.

112. *See id.* at 19; Goldstein, *supra* note 103, at 295–96.

113. IARC BENZENE STUDY, *supra* note 101, at 125.

not been the cause of death, and then argued that this cancer should be excluded from the total of excess mortality risks.[114] They also claimed that some included workers had been exposed to other carcinogenic substances, and therefore that the observed leukemias may not have been caused by benzene.[115] Indeed, Ott eliminated 53 workers including one leukemia victim from the cohort because they had been exposed to other "confounding exposures."[116] From the perspective of reliable science, it is obvious why a study should exclude ambiguous results in the presence of confounding variables. Given the limited scope of Ott's data, however, the effect of disregarding one leukemia that may have been caused by benzene was to reduce CAG's prediction of lifetime risks by nearly a third. This treatment is not clearly appropriate as a matter of social policy under conditions of continuing scientific uncertainty.

3. *The Infante Study and Rinsky Follow-Up*

The Infante study involved a retrospective cohort analysis of 748 workers employed in two rubber hydrochloride factories between 1940 and 1949. After determining the status of about 75% of the cohort between 1950 and 1975, Infante identified nine cases of leukemia mortality in comparison with an expected 1.25 leukemia deaths.[117] CAG interpreted the Infante data to find a relative risk of 7.20 resulting from the cohort's exposure to benzene,[118] and then estimated the unit risk factor as .014854 per ppm.[119]

Despite its reliance on the Infante study, CAG acknowledged several problems with its data:

> (1) [T]he authors essentially give no estimate of worker exposures except to say that the levels were less than the prevailing recommended occupational limits at the time various monitoring surveys were made; (2) . . . Air monitoring information at [one of the two plants] is almost non-existent, and therefore the exposure to half of the members of the cohort is almost completely unknown . . . (3)

114. The Agency properly rejected this criticism on the ground that the tumor might have been fatal if the subject had not died from other causes. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984).

115. *See id.*; Luken & Miller, *supra* note 40, at 1256.

116. *See* EPA Regulation of Benzene, 45 Fed. Reg. 23,478, 23,483 (1984); CAG BENZENE REPORT, *supra* note 96, at 18.

117. *See* CAG BENZENE REPORT, *supra* note 96, at 8–10.

118. *See id.* at 10. The Agency produced this estimate by dividing the observed leukemia incidence of 9 by the expected incidence of 1.25 cases.

119. *See id.* at 12–13.

> . . . [O]ver 400 workers known to be exposed to low benzene levels [may have been] deliberately excluded from the cohort.[120]

Because knowledge of historical exposure levels is necessary before a unit risk factor can be computed, CAG proceeded on the basis of one 1946 survey to "guess that the average exposure to all people in the plant before 1946 is probably not much more than 100 ppm, and not less than 15 ppm."[121] The Agency risk assessors also conceded: "Benzene levels were monitored after 1946 at various plant locations but they were all instantaneous samples and no reliable information is available about how many man-hours were spent at these locations or whether protective masks were worn."[122] CAG made "the assumption that the average worker exposure was the same as the prevailing recommended occupational limits."[123] Despite the problematical nature of these several assumptions and guesses, the Agency "felt that this study is the least flawed of the three utilized."[124]

In its 1984 benzene proceeding, EPA reviewed a follow-up study by Dr. Rinsky that extended coverage to 98% of the Infante cohort. This study adopted a strict cohort definition that excluded four leukemia cases which fell outside the job classifications or dates covered in the initial study,[125] but nonetheless found a statistically significant excess leukemia risk present among workers at both plants.[126] Rinsky and other scientists "for the most part" accepted Infante's contention that benzene exposures usually fell within the recommended limits.[127] After reviewing the available exposure data, however, the EPA official responsible for research on toxic effects emphasized:

> [I]n dealing with a low incidence phenomenon, it is difficult to utilize general industrial hygiene measurements as a means of typifying individual exposure. In both the Infante and Ott cohorts, less

120. *Id.* at 9.

121. *Id.* at 10. The CAG assessment recognized that no emissions monitoring records were available prior to 1946. *Id.*

122. *Id.* at 11.

123. *Id.* The Supreme Court plurality opinion in the benzene decision noted OSHA's finding that exposures in some plants may have considerably exceeded the 10 ppm PEL during the 1970's. *See* Industrial Union, Dep't, AFL-CIO v. American Petroleum Inst., 448 U.S. 607, 622 n.16 (1980). If periodic violations of a *legal* standard occurred during a decade when concern for toxic hazards was high, it is doubtful that most plants routinely complied with *voluntary* consensus standards in earlier decades.

124. CAG BENZENE REPORT, *supra* note 96, at 9; *see also* EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558, 23,560 (1984).

125. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,482 (1984); IARC BENZENE STUDY, *supra* note 101, at 125.

126. *See* IARC BENZENE STUDY, *supra* note 101, at 126.

127. *See id.* at 125-26; EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,482 (1984).

> than 1% of the total work force at risk developed leukemia. Accordingly, it is questionable whether the reported exposure levels are meaningful in estimating dose to the relatively few workers who developed the leukemia.[128]

Notwithstanding the problems with each epidemiological study, CAG used the geometric mean of the unit risk factors from its interpretation of the Aksoy, Ott, and Infante data to derive a predicted unit risk factor of .024074 per ppm for benzene.[129] This is the final, if magical, unit risk factor the Agency used to calculate the particular risks posed by different categories of benzene sources. In 1984 EPA reduced the CAG estimate seven percent to .022, reflecting minor revisions in interpretation of two studies,[130] but this revised unit risk factor was clearly based on the same data and risk-assessment assumptions used in the 1979 CAG analysis.

It is necessary to stress that epidemiologic evidence may be suitable for some quantitative purposes and not for others. With respect to the Infante study, for example, EPA noted that uncertainty about historical exposure levels is "irrelevant to the study's conclusion that exposed workers experienced a fivefold excess risk of leukemia over the general population."[131] After acknowledging the limitations in Ott's data, EPA similarly argued that this study "serves to reinforce the public health concerns regarding benzene exposure."[132] These comments indicate that the epidemiologic evidence was sufficient to prove benzene exposure induces leukemia, but EPA's characterizations do not demonstrate that the three studies could enable identification of the risks posed by specific exposure levels. In short, the conclusion that benzene is a carcinogen may be amply supported by data that are too indefinite to serve as the basis of a quantitative risk assessment. OSHA had considered the same epidemiological studies in detail and concluded that inadequate exposure data, small cohort size, and other methodological shortcomings precluded derivation of meaningful dose-response relationships for benzene risks.[133]

C. *Pollutant Concentrations and Dispersion Patterns*

EPA had to assess many issues other than the unit risk factor before producing final estimates of the leukemia hazards created by benzene discharges. Agency scientists employed emissions monitoring data and engi-

128. Goldstein, *supra* note 103, at 295-96.
129. *See* CAG BENZENE REPORT, *supra* note 96, at 20; Goldstein, *supra* note 103, at 296.
130. *See* EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,527 (1984).
131. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,482 (1984).
132. *Id.* at 23,483.
133. *See* Latin, *supra* note 5, at 361-64, 370 n.257; *supra* text accompanying notes 85-88.

neering estimates from specific plants, population distribution figures for areas around those pollution sources, meteorological data, and complex atmospheric dispersion models to predict the population exposed to varied concentrations of benzene.[134] Most of these essential variables are subject to considerable uncertainty.

There were, for example, fifty-five coke by-product recovery plants in 1984,[135] and EPA consultants identified more than twenty distinct emission sources in each plant that may produce varied discharges depending on individual plant designs.[136] It is not surprising the Agency acknowledged that "emissions and plant parameters often must be estimated rather than measured, particularly in determining the magnitude of fugitive emissions and where there are large numbers of sources."[137]

Estimation of ambient benzene concentrations depends on the atmospheric dispersion pattern associated with each plant's emissions, which is a function of localized meteorological conditions and terrain. Yet, sufficient information on these variables is difficult and expensive to obtain in practice. EPA conceded that "meteorological data often are not available at the plant site but only from distant weather stations that may not be representative of the meteorology of the plant vicinity."[138]

EPA noted that its best atmospheric dispersion model "is usually too resource intensive for modeling a large number of sources."[139] To simplify computational requirements, it relied on a dispersion model that excludes any pollution effects outside a twenty kilometer radius of each source "regardless of the estimated concentration at that point."[140] The computer model also assumed that the surrounding land is always flat. EPA recognized that "[f]or sources located in complex terrain, this assumption would tend to underestimate the maximum annual concentration,"[141] and might understate ambient levels by "several fold."[142]

134. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,492-93 (1984); *see also* EPA Draft Environmental Impact Statement: Benzene Emissions from Coke By-Product Recovery Plants—Background Information for Proposed Standards, Appendix E (May 1984) [hereinafter Benzene Draft EIS].

135. *See* Goldstein, *supra* note 103, at 299.

136. *See* EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,527-28 (1984).

137. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,493 (1984); *see also* EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558, 23,562 (1984).

138. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,493 (1984).

139. Benzene Draft EIS, *supra* note 134, at E-15.

140. *Id.*

141. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,493 (1984); *see also id.* at 23,492 nn.1 & 3.

142. Benzene Draft EIS, *supra* note 134, at E-15. The Agency also conceded that some of its other computerized exposure-assessment "assumptions are simplifications. People rarely live in the same place for 70 years; some move out and some move in. Nor do plants operate continuously for 70 years using the same equipment." EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,526 (1984).

The Agency attempted to offset these scientific deficiencies by claiming that other exposure-assessment factors may have been overestimated by an unspecified amount.[143] Given the great number of uncertain risk-assessment variables, EPA's frequent practice of presenting paired discussions of factors that may induce overestimation and underestimation of toxic risks can be used to rationalize any predictive judgment. Simply listing countervailing factors without any indication of their relative probability or importance does not reflect a systematic policy on how scientific uncertainties should be resolved in regulatory proceedings.

EPA's treatments of pollutant dispersion and population exposures were apparently shaped more by considerations of methodological convenience and resource constraints than by attempts to achieve the best possible science. The Agency's use of simplified analytical models cannot be criticized simply because they fail to capture every relevant facet of complex problems. Effective regulation may sometimes require agencies to adopt crude but administrable decisionmaking strategies that do not incorporate a high degree of scientific sophistication.[144] Yet, reliance on a range of debatable assumptions and analytical simplifications calls into question the overall scientific reliability of agency predictions. Complexity and uncertainty in estimating pollutant concentrations and exposed populations compound the imprecision of estimated unit risk factors for toxic substances.

D. *Animal Studies of Benzene Effects*

Reliable animal data on the carcinogenic effects of benzene were not available when OSHA and CAG conducted their initial risk assessments.[145] Because EPA's 1984 regulatory analysis depended heavily on the 1979 CAG assessment, the Agency did not rely on animal test data in its quantitative derivation of the unit risk factor. Instead, EPA merely cited the support of recent animal studies for its finding that benzene is a potent carcinogen.[146] In contrast, the California Air Resources Board and Department of Health Services (DHS) evaluated the results of several

143. *See* Benzene Draft EIS, *supra* note 134, at E-15; *see also* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,493 (1984); EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,526 (1984).

144. *See* Latin, *supra* note 11, at 1279-84, 1296-97, 1310-32.

145. *See* OSHA Occupational Exposure to Benzene, 43 Fed. Reg. 5918, 5932, 5946 (1978); CAG BENZENE REPORT, *supra* note 96, at 3.

146. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,484 (1984). CAG later conducted a quantitative risk assessment using the geometric mean of test results from four animal studies to derive a unit risk factor "almost identical" to its 1979 estimate based on human epidemiologic data. Goldstein, *supra* note 103, at 296, 298. Indeed, Dr. Goldstein noted that the closeness of the two estimates was "remarkable, perhaps even approaching witchcraft." *Id.* at 296.

animal studies in their quantitative assessment of benzene risks.[147] A brief discussion of the DHS analysis may be useful because most assessments of toxic risks depend heavily on extrapolation of human risks from animal test data.

DHS reviewed the results from a series of animal studies conducted by Dr. Maltoni and his associates and by the National Toxicology Program (NTP).[148] These studies unequivocally demonstrated that benzene "produces different types of tumors in different organs, and therefore it must be considered a multipotential carcinogen."[149] Based on cancer incidence data from the Maltoni and NTP experiments, DHS identified a series of possible human risks ranging between $20\text{–}340 \times 10^{-6}$ lifetime cancers per ppm of benzene.[150] However, the DHS report noted several important sources of uncertainty associated with basing risk estimates on the available animal studies.

Most of the cancers detected in the Maltoni and NTP studies affected the Zymbal glands of rats and the Zymbal and preputial glands of mice, but neither of these organ sites has any close biological analogue in human beings.[151] The DHS staff adopted protective IRLG and IARC positions that "it is appropriate to use the most sensitive species, sex, and tumor site in its assessments because there is often little correlation between tumor types or target organs between species which may in part be due to both physiologic differences between species and differences in the conditions of the bioassay and actual human exposure."[152] This treatment may be appropriate to establish that a substance is carcinogenic, but it is much more problematic whether animal test data derived from different organs and tumor types can be used to derive a reliable dose-response curve and unit risk factor applicable to human beings.

Animal tests typically entail high doses administered over short periods of time, but human hazards from toxic substances are usually characterized by long-term exposures to relatively low doses.[153] Application of high dosages in animal studies presents several problems. For example, the DHS report noted:

147. *See* DHS BENZENE REPORT, *supra* note 39, at 34–126. The state agencies conducted an extended assessment of benzene risks because the majority of atmospheric benzene emissions originate from automobile exhausts and gas station operations, which are sources of special concern in Southern California.

148. *See id.* at 84–99.

149. *Id.* at 89.

150. *Id.* at 99.

151. *Id.* at 89, 91; Goldstein, *supra* note 103, at 296.

152. DHS BENZENE REPORT, *supra* note 39, at 65.

153. *See* Latin, *supra* note 5, at 377–79.

> Due to the high cost of animal bioassays, studies are conducted with relatively few animals (usually 50) per dose group. This results in a very low statistical power to detect small increases in disease rates. To compensate for this, high dose schedules are used. This assumes that chemically induced carcinogenic responses at high doses will also result in similar responses at low doses.[154]

In comparison to small test groups, a cohort of about 10,000 animals would be required to detect a statistically significant risk of one in 100 at realistic human dosage levels.[155] Yet, high doses may affect carcinogenesis processes and may distort low-dose-response calculations. DHS noted that animals in both the NTP and Maltoni studies had "significant noncarcinogenic adverse toxicological responses" including "dose-related increased mortalities and dose-related weight losses."[156] These "chronic toxicological insults from high doses of benzene" may have caused or contributed to the observed incidence of cancers.[157] DHS provisionally rejected this indirect causal mechanism as unproven,[158] but the Agency noted that the level of scientific knowledge is weak in this area.[159]

DHS observed that test results could be affected by "animals which die during the 2-year course of the experiment either due to benzene toxicity or due to natural causes, and thus are not available to develop cancer."[160] In other words, high test doses could cause nonmalignant mortalities and correspondingly reduce the apparent incidence of cancers in the study. DHS indicated that an adjustment for premature mortality in test subjects might produce up to a fourfold increase in the predicted unit risk factor, but the Agency did not actually make this correction in its final assessment of benzene risks.[161]

Because animal test data stem from high-dose exposures, a risk assessment must employ an extrapolative theory to predict low-dose effects and an interspecies scaling factor that relates animal doses to human exposures. DHS considered numerous extrapolative theories that have been proposed in the scientific literature and eventually followed EPA's choice

154. DHS BENZENE REPORT, *supra* note 39, at 66.

155. *See* Ames, *Identifying Environmental Chemicals Causing Mutations and Cancer*, 204 SCIENCE 587, 589 (1979); Leape, *supra* note 6, at 93-94.

156. DHS BENZENE REPORT, *supra* note 39, at 98.

157. *Id.*

158. *See id.* at 52, 54-55, 67, 98-99.

159. *See id.* at A1-A5.

160. *Id.* at 100.

161. *See id.* 100-101. The DHS report's failure to adjust its final risk estimates for animals that died of nonmalignant causes during the studies may explain the discrepancy between its "high" estimate in one passage of 340×10^{-6}/ppb, *id.* at 99, and another "high" estimate of 170×10^{-6}/ppb, *id.* at 110. These estimates differ by a factor of two, and DHS noted that an adjustment for competing causes of death could lead to higher risk level estimates. *Id.* at 100.

of a multistage model.[162] The state Agency, however, noted: "It is not possible to validate any of these models in the low dose range with either animal or epidemiologic data so selection of a particular model is somewhat arbitrary."[163] DHS also observed that several different interspecies scaling factors could be adopted before it selected the surface-area criterion.[164] DHS concluded that "the mg/surface area-day factor is most appropriate because it falls near the middle of the range of measures that have been proposed."[165] The Agency did not provide any scientific or social policy explanation why a mid-range position is preferable to a conservative one. In justifying its choice of a protective extrapolative theory for low-dose exposures, DHS contended that "the more conservative of equally reasonable elements should constitute the basis for regulation."[166] Yet DHS did not adopt an equivalent approach in its choice of an interspecies scaling factor, nor explain the inconsistency in its treatments of uncertainty.

This discussion only suggests the many scientific uncertainties that preclude reliable derivation of human carcinogenic risks from animal test results. In the recent *Science* symposium on risk assessment, three scientists observed:

> Quantitative extrapolation from rodents to humans, particularly at low doses, is guesswork that we have no way of validating. It is guesswork because of lack of knowledge in at least six major areas: (i) the basic mechanisms of carcinogenicity; (ii) the relation of cancer, aging, and lifespan; (iii) the timing and order of the steps in the carcinogenic process that are being accelerated; (iv) species differences in metabolism and pharmacokinetics; (v) species differences in anticarcinogens and other defenses; and (vi) human heterogeneity—for example, pigmentation affects susceptibility to skin cancer from ultraviolet light. These sources of uncertainty are so numerous, and so substantial, that only empirical data will resolve them, and little of this is available.[167]

The symposium article concluded "it is not scientifically credible to use results from rodent tests done at [high doses] to directly estimate human

162. *See id.* at 68-80.

163. *Id.* at 69.

164. *See id.* at 68, 106; *supra* text accompanying notes 67-70.

165. DHS BENZENE REPORT, *supra* note 39, at 68.

166. CALIFORNIA AIR RESOURCES BD. & DEP'T OF HEALTH SERVS., REPORT TO THE SCIENTIFIC REVIEW PANEL ON BENZENE: OVERVIEW AND RECOMMENDATION 8 (Nov. 1984) [hereinafter DHS BENZENE OVERVIEW].

167. Ames, Magaw & Gold, *Ranking Possible Carcinogenic Hazards*, 236 SCIENCE 271, 275 (1987) (citations omitted).

risks at low doses."[168] Yet, this is precisely the function of the risk-assessment process, and the DHS benzene analysis relied heavily on extrapolating human risks from animal test results.

DHS also reviewed more than a dozen epidemiological studies of benzene effects including the three on which EPA relied.[169] The state report noted the value of epidemiologic data for establishing that benzene is a carcinogen, but then identified many weaknesses in the major studies.[170] DHS nonetheless followed CAG's risk-assessment assumptions to calculate a best-guess unit risk factor of 48×10^{-6} leukemia cases based on the Rinsky reinterpretation of Infante's study, which was roughly double the 22×10^{-6} EPA estimate derived from the geometric mean of the three studies.[171] The DHS staff, however, chose to describe leukemia hazards as a series of risk estimates ranging between EPA's low prediction and a high estimate of 170×10^{-6} per part per billion (ppb) arising from extrapolation of animal test results.[172] In other words, DHS presented the EPA estimate, its own estimate based on the Rinsky study, and several other predictions derived from animal studies to identify a range of possible risk values that varied by a factor of about eight.[173]

168. *Id.* at 277.

169. *See* DHS BENZENE REPORT, *supra* note 39, at 42-44.

170. DHS listed the following problems: "Exposure levels and exposure periods are poorly documented, mortality rather than incidence is reported, the number of exposed individuals tends to be small, appropriate control groups are not always used, results are only directly applicable to white employed males—effects in women and children have not been sufficiently studied, and few confounding factors are controlled for." *Id.* at 110. DHS also cited "the long latency period for the development of human cancers, [and] the difficulty of identifying a large appropriate study population" as reasons why "epidemiological studies are of limited usefulness as a means of carcinogen identification." *Id.* at 33. After summarizing criticisms of the Aksoy, Ott, Infante, and Rinsky studies, *see id.* at 57, 62-63, the DHS report noted:

> Much of the criticism of the CAG risk assessment related directly to the quality of the benzene exposure data; specifically, the exposure period and the exposure levels. For example, the exposure period in the Rinsky study is taken to be 35 years, an estimate some consider to be longer than the actual exposure time. The use of this exposure period has the effect of increasing the cumulative exposure dose and hence yields a slope that is lower than the actual. Further, since no routine monitoring of the work place was performed during the periods workers in this study were exposed, much controversy has arisen as to the level of benzene workers were exposed to. Here it is argued that [the] exposure level used for the assessment substantially underestimated the true exposure level thereby resulting in an overestimate of the slope.

Id. at 63.

171. *See id.* at 57-59, 114-15.

172. *See id.* at 108-11. In the same vein, DHS stated that its analysis of the Rinsky data produced a risk estimate between 32×10^{-6} and 120×10^{-6}, with 48×10^{-6}/ppb as the most likely unit risk factor computed directly from Rinsky's study. *Id.* at 58-59, 114-15. The disparity of risk estimates, unlike the one obtained by DHS from animal studies, was a function of statistical confidence levels rather than a difference in risk-assessment assumptions. *See id.* at 59.

173. *See id.* at 108 (Figure 3). The Air Resources Board estimated that people in the southern region of California are exposed to average concentrations of 4.6 ppb, and the Board then used the range of DHS risk estimates to predict that benzene emissions would cause between 101 and 780 leukemia cases per million lifetime exposures. *See* DHS BENZENE OVERVIEW, *supra* note 166, at 8.

As in EPA's carcinogen guidelines and benzene proceedings, the DHS benzene assessment mixed conservative, best-current-guess, mid-range, and methodological-convenience positions on various issues without any thematic coherence. If DHS had selected a series of more conservative treatments—for example, by adjusting for premature animal subject mortality and by using the mg/kg-lifetime scaling factor and linear one-hit extrapolative theory—the revised benzene risk estimate could have been an order of magnitude or more greater than the predictions presented in the state report. On the other hand, adoption of less conservative assessment criteria might have led to an estimate of only negligible risks from ambient levels of benzene. Because critical uncertainties cannot be resolved exclusively through reliable scientific judgments, the ultimate regulatory decisions must explicitly or implicitly incorporate policy judgments.

E. *Inconsistent Treatments of Uncertainty*

In its 1984 benzene proceedings, EPA multiplied the ambient concentration and population estimates by the calculated unit risk factor to identify predicted leukemia hazards from five types of industrial sources.

TABLE 3

Estimated Leukemia Risks From Industrial Sources[174]

Category of Benzene Discharger	Level of Emissions (Mg/year)	Maximum Lifetime Risk/Most Exposed Population Group	Annual Leukemia Incidence For Most Exposed Population
Coke By-Products	24,100	6.4×10^{-3}	2.2
Fugitive Emissions	7,900	1.5×10^{-3}	0.45
Benzene Storage Vessels	620	3.6×10^{-5}	0.043
Maleic Anhydride	960	7.6×10^{-5}	0.029
Ethylbenzene/ Styrene	210	1.4×10^{-4}	0.0057

174. The figures for all source categories except coke by-product recovery plants are presented in Table 1 of EPA's explanation for its selective regulation of benzene dischargers. See EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,492 (1984). The other estimate is from EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,525 (1984).

In light of the uncertainties underlying toxic risk assessment, EPA's estimates of .0057 cases per year—one leukemia every other century—from ethylbenzene plants and of fewer than three cases per century from maleic anhydride discharges are surely artifacts of the Agency's arithmetic rather than scientifically credible figures. Nevertheless, these risk-assessment estimates formed the foundation for EPA's decision to promulgate standards for the first two categories while withdrawing proposed controls on other industrial sources. The fundamental question is whether EPA's predictions are reasonably accurate, perhaps within one or two orders of magnitude, but scientists cannot answer this question dispositively given the imperfect state of current knowledge. Regulators must consequently decide how they should allocate the legal and social burdens associated with continuing scientific uncertainty.

In published explanations of its benzene decisions, EPA claimed its analyses "represent plausible, *if conservative*, estimates of the magnitude of the actual human cancer risk posed by benzene emitted from the source categories evaluated."[175] The Agency was, however, forced to retreat from this claim in several instances. The Natural Resources Defense Council (NRDC), for example, challenged CAG's derivation of the unit risk factor through equal weighting of the three epidemiological studies. NRDC argued that primary weight should have been given to the study yielding the highest risk estimate.[176] As noted above, CAG not only used the geometric mean rather than the highest risk prediction among the three studies, but also took the geometric mean of historical exposure estimates in computing the unit risk factor for each study.[177] Neither averaging treatment explicitly provides "an ample margin of safety" under conditions of uncertainty, as the statute mandates.[178] EPA also acknowledged that some of its benzene exposure predictions "may overestimate or underestimate actual emissions from individual sources," and these estimates "do not reflect a systematic conservatism."[179] In one document EPA candidly stated its position that "[t]he choice of the most conservative assumption in each case would result in estimates unreasonably biased in the direction of overestimation."[180] The Agency then maintained that its "risk assessment

175. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,493 (1984) (emphasis added). *See also* EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558, 23,560 (1984); EPA Benzene Emissions from Maleic Anhydride Plants, Ethylbenzene/Styrene Plants, and Benzene Storage Vessels; Proposed Withdrawal of Proposed Standards, 49 Fed. Reg. 8386, 8388 (1984).

176. EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558, 23,560 (1984).

177. *See supra* text accompanying notes 98-100, 121-24.

178. *See* 42 U.S.C. §7112(b)(1)(8) (1982).

179. EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558, 23,562 (1984).

180. *Id.* at 23,560.

provides plausible, *if not conservative*, estimates of the benzene health risks."[181]

The analysis in this Part does not demonstrate that EPA must always adopt the most protective possible treatment, nor that any particular assessment practice in its benzene proceedings was necessarily improper. Rather, under any test of administrative rationality, EPA's responses to scientific uncertainty were neither consistent nor cogently explained. The Agency selected conservative positions on some important issues, such as its adoption of a nonthreshold extrapolation theory.[182] In other instances EPA relied on geometric means or mid-range positions, such as its treatment of the sparse data on past exposure levels. On some issues, such as pollutant dispersion modeling, the Agency employed methodologically convenient but relatively unrealistic treatments. And in some contexts, EPA adopted risk-assessment practices that are likely to result in underestimation of important variables. For example, CAG extrapolated from occupational studies of "generally healthy, white males" to the "general population for whom susceptibility to a carcinogenic insult could differ."[183] Yet, EPA "is uncertain whether the unit risk factor can be accurately applied to the general population, which includes men, women, children, nonwhites, the aged, and the unhealthy."[184] There was no unifying logic in the Agency's treatments of different scientific uncertainties, and EPA has never provided a coherent public explanation of its disparate practices. Indeed, the regulatory record provides little indication that Agency risk assessors recognized the diversity or social implications of their responses to uncertainty. It is even less likely that EPA risk managers were aware of the inconsistent treatments of uncertainty, which were obscured by the presentation of precise quantitative risk estimates.

I have not described the technical issues in EPA's benzene proceedings only to demonstrate that scientific uncertainty pervades the risk-assessment process, although that observation is essential for any analysis of toxic substances regulation. My discussions with people in this field suggest that few environmental or administrative law experts recognize (1) the extent to which agencies adopt inconsistent treatments of diverse uncertainties, (2) the tendency of risk assessors to trivialize uncertainty by listing potentially offsetting factors without evaluation of their relative significance, (3) the degree to which implicit social policy positions shape

181. *Id.* (emphasis added).

182. *See* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,484-88 (1984).

183. *Id.* at 23,493.

184. Benzene Draft EIS, *supra* note 134, at E-16. The Agency also observed that the exposed population is a "large, diverse, and genetically heterogeneous group" and that "genetic variability to carcinogenesis is well documented." EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,485 (1984).

risk assessments that are ostensibly scientific in nature, and (4) the failure of agencies to explain fully the bases for their selection of controversial risk-assessment positions. These types of findings can be made only through careful examination of actual regulatory proceedings, and yet an understanding of such "details" is crucial for development of effective risk-assessment strategies.

III. Ramifications of EPA's Emphasis on "Good Science"

The risk-assessment practices described above indicate that EPA's attempt in its carcinogen guidelines and benzene proceedings to base risk estimates on "the most scientifically appropriate interpretation"[185] entails several controversial social ramifications:

A. *Trade-offs Between the Pursuit of "Good Science" and Effective Protection Under Uncertainty*

Although current guidelines embody some conservative risk-assessment principles, the individualized "weight of evidence" approach coupled with agency attempts to tailor all analyses in light of changing scientific knowledge will often reduce the degree of protection previously afforded by the IRLG guidelines. Few if any of the revised treatments in the guidelines have achieved general scientific acceptance, and EPA does not contend that most uncertainties can be resolved with reasonable scientific assurance. Given the imperfect state of the risk-assessment art, regulators must decide how much potential but uncertain public protection should be traded for some potential but uncertain improvement in the accuracy of scientific judgments that EPA clearly recognizes are far from reliable. The present guidelines assume that every tentative step, however provisional, in the direction of "good science" is warranted regardless of its possible effect on the scope of protection. The wisdom of this presumption is surely a public policy issue rather than a purely scientific question.

The new guidelines reflect a relative shift in EPA's emphasis on two recurring questions in toxic substances regulation: Is there sufficient reliable evidence that a chemical produces "toxic" effects at high or unknown past exposure levels, and is there enough evidence to derive reliable quantitative risk-assessments at specific exposure levels.[186] If the Agency delays regulation until the "weight of evidence" enables predictions about specific dose-response relationships, as the guidelines presume, then EPA may

185. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992 (1986).

186. For a discussion of this important distinction in the two kinds of evidence on toxicity, see Latin, *supra* note 5, at 344–45.

allow years of continued exposures to a known toxic substance because the precise level of toxicity cannot be reliably estimated. The Agency decision to wait until regulators can meet the particularized evidentiary requirements of the guidelines is equally a decision to stress scientific validity rather than safety after an indeterminate toxic hazard has been qualitatively identified. This preference is neither inevitable nor consistent with past practices.

As one illustration of how a requirement for "good science" in regulatory determinations can affect the scope of public protection, the Clean Water Act initially provided that EPA must control toxic pollutants based on their degree of toxicity.[187] This harm-based regulatory strategy presupposed that EPA could produce particularized assessments of the hazards created by specific substances. After the Agency's failure to issue any toxic water pollutant standards was challenged in litigation, EPA adopted a "technology-based" approach in which it imposed strict standards based on qualitative proof that a substance is "toxic" and that controls are technologically and economically feasible.[188] The EPA Assistant Administrator for Water and Hazardous Materials testified before Congress that the original approach was "technically impractical" because the Agency could not "demonstrate the cause and effect relationship between pollutants and public health."[189] Administrator Costle similarly testified in 1977 that "experience with the alternative approaches . . . leave[s] us firmly convinced that for the bulk of known or suspected toxics of concern, technology-based standards established on an industry-by-industry basis are by far the most feasible to implement and administer."[190] In short, EPA adopted the technology-based regulatory approach because the Agency lacked the information necessary to perform quantitative risk assessments and because it decided that protective regulatory action was necessary despite scientific uncertainty.[191] The carcinogen guidelines, in contrast, require precisely the kind of individualized evidence that EPA had previously found difficult to obtain, and in recent years the Agency has promulgated few harm-based standards for toxic water pollutants.

The current EPA guidelines claim that the "impetus for this revision is

187. *See* 33 U.S.C. §1317(a) (1982).

188. *See* W. RODGERS, ENVIRONMENTAL LAW 486–87 (1977); Latin, *supra* note 11, at 1307–09.

189. EPA Focuses on 1983 BAT Requirements to Control Most Toxic Pollutants, 8 ENV'T REP. (BNA) 476 (1977) (reporting the testimony of Thomas Jorling).

190. HOUSE COMM. ON PUB. WORKS AND TRANSP., 95TH CONG., 1ST SESS., IMPLEMENTATION OF THE FEDERAL WATER POLLUTION CONTROL ACT: SUMMARY OF HEARINGS ON THE REGULATION AND MONITORING OF TOXIC AND HAZARDOUS CHEMICALS UNDER THE FEDERAL WATER POLLUTION CONTROL ACT (Pub.L. No. 92-500, 86 Stat. 816 (1972)) 26 (Comm. Print 1977).

191. *See* Latin, *supra* note 11, at 1308–09. I believe EPA has promulgated more toxic substances standards in the past decade under this one technology-based program than it has under all of its programs that require quantitative risk assessments based on "good science."

the need to incorporate . . . the concepts and approaches to carcinogen risk assessment that have been developed during the last ten years."[192] Yet, the guidelines never specify which scientific advances now enable reliable risk assessments, and they acknowledge the existence of the same uncertainties that previously impeded regulation. Of special concern, the guidelines do not indicate what regulatory actions are appropriate during the often lengthy period between the time a substance has been identified qualitatively as "toxic" and the time quantitative risk estimates become practicable. To the extent administrators are now required to support regulation of carcinogens with the kind of "weight of evidence" assessments envisioned in the guidelines, this position clearly places the burden of scientific uncertainty on exposed populations.

B. *Effects on Agency Behavior*

The pursuit of "good science" based on individualized circumstances is likely to increase the decisionmaking costs and time requirements associated with the risk-assessment process. With respect to animal tests, for example, the guidelines state that the "weight of evidence" for potential human hazards rises "with the increase in number of animal species, strains, sexes, and number of experiments and doses showing a carcinogenic response."[193] With respect to data from epidemiological studies, the guidelines similarly observe that the "weight of evidence increases rapidly with the number of adequate studies that show comparable results on populations exposed to the same agent under different conditions."[194] Both types of studies are expensive, may take years to complete, and are frequently inconclusive. The carcinogen guidelines, however, never address EPA budgetary restrictions or the time-lag, with accompanying irreversible health effects, that may occur while regulators wait for sufficient data to make reliable scientific judgments.

A more subtle ramification is that the guidelines invite Agency officials to evaluate their own performance, and that of their subordinates, in terms of scientific competency rather than regulatory competency. If the primary decisional criteria is whether regulators select the "most scientifically appropriate interpretation to assess risk," officials may be reluctant to choose speculative treatments that increase public safety under conditions of uncertainty but cannot be identified as the most plausible scientific

192. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,993 (1986).

193. *Id.* at 33,994. "In general, although a single study may be indicative of a cause-effect relationship, confidence in inferring a causal association is increased when several independent studies are concordant in showing the association . . . ." *Id.* at 33,999.

194. *Id.* at 33,995.

theories among a constellation of competing hypotheses. Moreover, the majority of interveners in regulatory proceedings are sponsored by affected industries or trade associations,[195] which means the scientific performance of agency officials will regularly be monitored and challenged by industry scientists who advocate less conservative risk assessment practices. Agency bureaucrats, like other people, are sensitive to criticism and may deliberately or subconsciously seek to placate persistent critics.[196]

C. *Increased Opportunities for Obstructive Behavior by Affected Parties*

Even if Agency risk assessors are assumed to be motivated solely by a desire to conduct the best possible scientific analyses based on the available evidence, a comparable assumption cannot be applied to the goals of interveners who espouse conflicting private interests.[197] The primary incentive of industry representatives is to minimize regulatory costs, not to promote good science. The primary interest of environmentalist interveners is to minimize health and ecological risks irrespective of regulatory costs, not to promote good science. The "weight of evidence" approach embodied in the carcinogen guidelines allows parties in each proceeding to make any conceivable scientific argument—and some inconceivable ones if past practices are any guide—that may affect agency decisionmaking directly through the force of debatable scientific arguments or indirectly through increased delays and costs.

D. *Increased Opportunities for Abuse of Discretion by Agency Decisionmakers*

Emphasis on individualized "weight of evidence" judgments may enable regulators to make ideologically motivated decisions under the guise that they represent "good science." In 1982, for example, the EPA Assistant Administrator for Pesticides and Toxic Substances, John Todhunter, concluded that formaldehyde poses only a low carcinogenic risk which need not be regulated under the Toxic Substances Control Act.[198] This decision ostensibly was predicated on the Agency's risk assessment, not on regulatory cost considerations or political values incorporated in the risk-management process. Todhunter's formaldehyde risk assessment, however, incorporated many questionable analytical assumptions. He presumed that a safe threshold level exists for low exposures, that only body-site specific tumors should be counted in the test results, that positive animal

195. *See, e.g., id.* at 34,001; EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,480 (1984).
196. *See* Bayley, *Memoirs of a Fox*, 2 ENV'T AFF. 332 (1972-73).
197. *See* Latin, *supra* note 11, at 1292-97.
198. *See* Ashford, Ryan & Caldart, *supra* note 32, at 324-43, 346-54.

tests do not fairly indicate human hazards, and that vulnerable people will protect themselves because high exposures may cause unpleasant irritant effects.[199] EPA adopted these assumptions on an ad hoc basis despite their inconsistency with previous Agency practices and with risk-assessment principles widely held in the scientific community. Indeed, the contemporaneous scientific literature sharply criticized Todhunter's analytical positions and conclusions.[200] After Todhunter left office, the Agency reopened the formaldehyde issue and decided in 1984 that two categories of emissions sources should be regulated.[201]

Although the analytical principles in the EPA's carcinogen guidelines and 1984 benzene assessment frequently conflict with the practices selected by Todhunter, nothing in the guidelines would prevent an Agency official from intentionally using debatable outcome-oriented assumptions whenever he or she asserts that those risk-assessment treatments are warranted by the weight of scientific evidence. For example, the guidelines provide: "Evidence indicating that high exposures alter tumor responses by indirect mechanisms that may be unrelated to effects at lower exposures should be dealt with on an individual basis."[202] This provision apparently would allow Agency decisionmakers to find that a safe threshold exposure level exists for a particular carcinogenic substance. The public policy problem with this degree of quasi-scientific discretion is that regulatory judgments expressed as individualized findings of "good science" are likely to be less visible and more immune from effective judicial or legislative review than decisions clearly based on economic concerns or controversial political values.

E. *Susceptibility to Intrusive Judicial Review*

Agency contentions that toxic controls are grounded on "good science" may increase the vulnerability of regulations to hostile judicial review. The CPSC, for example, tried to regulate urea-formaldehyde foam insulation on the basis of one experiment in which more than 40% of the animals contracted cancer within 24 months.[203] This finding showed an unusually high degree of carcinogenic potency in comparison with the

199. *See id.* at 330–32; Latin, *supra* note 11, at 1327–28.

200. *See* Ashford, Ryan & Caldart, *Law and Science Policy in Federal Regulation of Formaldehyde*, 222 SCIENCE 894 (1983); Hileman, *Formaldehyde; How Did EPA Develop Its Formaldehyde Policy?*, 16 ENVTL. SCI. & TECH. 543 (1982); Marshall, *EPA's High-Risk Carcinogen Policy*, 218 SCIENCE 975 (1982); Perera & Petito, *Formaldehyde: A Question of Cancer Policy?*, 216 SCIENCE 1287 (1982).

201. *See* EPA Formaldehyde; Determination of Significant Risk, 49 Fed. Reg. 21,870 (1984).

202. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,995 (1986).

203. *See* Gulf S. Insulation v. Consumer Prod. Safety Comm'n, 701 F.2d 1137, 1141, 1146 (5th Cir. 1983); Ashford, Ryan & Caldart, *supra* note 32, at 317–20.

animal data on other toxics in widespread use.[204] In *Gulf South Insulation v. Consumer Product Safety Commission*, the Court of Appeals for the Fifth Circuit overturned the formaldehyde regulation because the judicial panel decided "it is *not good science* to rely on a single experiment, particularly one involving only 240 subjects, to make precise estimates of cancer risk . . . . *To make precise estimates, precise data are required.*"[205] The opinion provided no intimation of how much precision is required for risk estimates nor how much "precise data" are necessary to constitute substantial evidence in support of regulatory judgments. The Fifth Circuit judges, however, apparently were prepared to make this decision themselves rather than to defer to Agency determinations.

The court's opinion reflects insensitivity to the protective goals of the organic regulatory legislation[206] and a fundamental misunderstanding of the limited evidence on which most risk assessments of carcinogens are based. Urea-formaldehyde foam insulation was a relatively new product, which precludes the acquisition of human epidemiologic data given long latency periods for many forms of cancer. With respect to the animal data, 240 subjects is a relatively large cohort in comparison with most experiments.[207] Moreover, virtually all regulatory discussions of toxic hazards agree that test results from the most sensitive species and exposure conditions should receive special weight in risk assessments of carcinogens.[208] Thus, findings of lower potency in subsequent studies would not necessarily negate the significance of the initial finding of high toxic potency. Many, if not most, quantitative risk assessments based on animal test data have relied on findings from one experiment or one series of related tests conducted by a single group of experimenters. The Court of Appeals opinion seems to assume that valid science requires agencies to average the results of several positive tests before developing a quantitative risk assessment or finding that a substance causes cancer in humans. This judicial conclusion is not generally accepted in the scientific community nor warranted from the viewpoint of good regulation.[209]

204. For discussions of the animal test data on benzene, see DHS BENZENE REPORT, *supra* note 39, at 84–93; *supra* text accompanying notes 145–66. CPSC had also attempted to regulate benzene on the basis of very limited data. *See* Luken & Miller, *supra* note 40, at 1255.

205. *Gulf S. Insulation*, 701 F.2d at 1146 (emphasis added).

206. The court did not consider the social consequences of allowing the toxic hazard to continue while the Agency tries to accumulate precise data on formaldehyde cancer risks, nor whether this allocation of the burden of uncertainty is compatible with the principles and priorities in the organic act. *See* Latin, *supra* note 11, at 1327–29 & n.305.

207. For a discussion of the animal cohorts used in studies of benzene toxicity, see DHS BENZENE REPORT, *supra* note 39, at 84–99.

208. *See, e.g.*, EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986); DHS BENZENE REPORT, *supra* note 39, at 65.

209. EPA's current carcinogen guidelines observe that a single study could demonstrate a sufficient cause-effect relationship, although the Agency would prefer results from additional sources. *See*

In *Industrial Union Department, AFL-CIO v. American Petroleum Institute*,[210] the Supreme Court invalidated OSHA's 1978 benzene regulation on the ground that the Agency had failed to prove its standard was "reasonably necessary" to prevent a "significant" risk of harm at prevailing exposure levels.[211] Compliance with this judicially-imposed "significance" test presupposes quantitative risk assessments at specific exposure levels, despite OSHA's contention that it could not produce reliable dose-response estimates for carcinogens.[212] Notwithstanding its many analytical deficiencies,[213] the Court's benzene decision has been followed by courts in a variety of toxic regulation contexts[214] and has induced federal agencies to conclude that they must provide quantitative risk estimates even if they lack confidence in the resulting judgments.[215] In response to the Court's opinion, OSHA indicated that it would determine whether potential carcinogens pose "significant" risks at specific exposure levels after individualized analyses of the scientific evidence.[216] Under the Reagan Administration, however, OSHA has not reduced the permissible discharge limit for any industrial benzene source below the level initially set in 1969.[217]

Contemporaneous with its 1978 benzene proceedings, OSHA was developing a generic cancer policy intended to preclude individualized analyses of speculative scientific theories and supporting data that the Agency had previously considered and rejected.[218] The plurality opinion

EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,999 (1986); *cf. supra* note 62 (data from most sensitive species should be used "in absence of evidence to the contrary").

210. 448 U.S. 607 (1980).

211. *Id.* at 639-40, 652-59.

212. *See* Latin, *supra* note 5, at 344-45. The Court confused the use of quantitative evidence to show excess deaths from all levels of past benzene exposure with the use of that evidence to establish the risks created by particular emission levels. When past exposure levels are unknown or are much higher than current levels, epidemiologic data may document carcinogenic hazards from past exposure without providing any insight on the risks arising from current discharge levels. I described this as a distinction between "aggregate" and "disaggregated" risks. *See id.* at 346, 385-86.

213. There is no need to repeat here my previous criticisms of the plurality opinion's inept treatment of the scientific evidence and its confused interpretation of the toxic substances provision in the OSH Act. *See* Latin, *supra* note 5; *see also* Sullivan, *The Benzene Decision: A Contribution to Regulatory Confusion*, 33 ADMIN. L. REV. 351 (1981).

214. *See, e.g.*, American Textile Mfrs. Inst. v. Donovan, 452 U.S. 490, 505 n.25 (1981); Asbestos Information Ass'n v. OSHA, 727 F.2d 415, 424-26 (5th Cir. 1984); Gulf S. Insulation v. Consumer Prod. Safety Comm'n, 701 F.2d 1137 (5th Cir. 1983); United Steelworkers v. Marshall, 647 F.2d 1189 (D.C. Cir. 1980), *cert. denied*, 453 U.S. 913 (1981); Texas Indep. Ginners Ass'n v. Marshall, 630 F.2d 398 (5th Cir. 1980).

215. *See* EPA Withdrawal of Proposed Standards, 49 Fed. Reg. 23,558, 23,559 (1984); Cross, *supra* note 25, at 12-43; Preuss & White, *supra* note 43, at 333 (Supreme Court's benzene decision had "probably the greatest impact" in increasing agency reliance on quantitative risk assessment).

216. *See* OSHA Indentification, Classification and Regulation of Potential Occupational Carcinogens; Proposed Amendments, 46 Fed. Reg. 7402 (1981).

217. Because the existing 10 ppm benzene standard was established to prevent nonmalignant blood diseases, *see supra* note 85, the upshot is that OSHA has never imposed regulatory restrictions based on the carcinogenic effects of widespread benzene exposures.

218. *See* OSHA Identification, Classification and Regulation of Toxic Substances Posing a Poten-

in *Industrial Union* presumed that OSHA had impermissibly relied on this generic policy in its benzene proceeding.[219] The Court then intimated that the Agency must always consider potentially relevant scientific evidence on an individualized basis.[220] This judicial treatment was partly grounded on the erroneous assumption that OSHA had been able in other toxics proceedings to obtain sufficient evidence on cancer risks at specific exposure levels to meet the "significance" test.[221] The Court never considered the impacts of its requirement for individualized risk assessments on agency decisionmaking costs and regulatory delays. Unfortunately, the treatment mandated by the plurality opinion invites affected parties to submit the same kinds of speculative theories and data in every hearing on toxic hazards. This is precisely what has occurred in subsequent proceedings on benzene conducted by OSHA, EPA, and other agencies.[222]

Unrealistic judicial requirements for comprehensive agency assessments of all potentially relevant factors and for a high degree of scientific precision have substantially emasculated environmental control programs in the past decade.[223] Yet, EPA's current "good science" orientation exacerbates this problem. Regulated industries and other interveners invariably can challenge the scientific bases of carcinogen risk assessments because uncertainty is pervasive and agency officials must adopt many debatable procedures in response to resource constraints and limited data. If regulators

tial Occupational Carcinogenic Risk, 42 Fed. Reg. 54,146 (1977). This generic policy was formally adopted at 45 Fed. Reg. 5002 (1980), but has been almost entirely ignored by the Agency under the current administration.

219. *See* Latin, *supra* note 5, at 364-67.

220. OSHA itself interpreted the Supreme Court's decision in this manner, and amended the generic cancer policy to indicate that it would make individualized determinations of carcinogenic risks. OSHA did not, however, explain how it would implement this policy in practice. *See* OSHA Indentification, Classification and Regulation of Potential Occupational Carcinogens; Conforming Deletions, 46 Fed. Reg. 4889 (1981).

221. Justice Stevens cited several regulatory proceedings where he thought OSHA had produced sufficient individualized evidence on cancer risks, *see* Industrial Union Dep't, AFL-CIO v. American Petroleum Inst., 448 U.S. 607, 656-57 & n.64 (1980), but he clearly misunderstood and mischaracterized the scientific evidence submitted on those other carcinogens. *See* Latin, *supra* note 5, at 369-80, 384-88.

222. In proceedings on benzene risk levels held after the Supreme Court's decision, industry representatives made many of the same arguments and submitted the same questionable data that OSHA had found unpersuasive in its 1978 benzene regulatory hearing and generic cancer policy. *Compare* Latin, *supra* note 5, at 361-64, 367-71, 378-79 *with* EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,483 (1984) *and* EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,527 (1984).

223. *See* Huber, *supra* note 74; Latin, *supra* note 5; Stewart, *supra* note 11, at 1274-75, 1338-40. Professor Jerry Mashaw draws similar conclusions in a study on the effects of judicial review on automobile safety regulation. *See* Mashaw & Harfst, *Regulation and Legal Culture: The Case of Motor Vehicle Safety*, 4 YALE J. ON REG. 257, 312-313 (1987). I find persuasive Mashaw's conclusion that intrusive judicial review has paralyzed rulemaking programs in complex technical areas because appellate courts focus on the adequacy of agency decisionmaking, which inevitably appears imperfect in retrospect, rather than on achievement of the protective goals of the regulatory statutes.

explicitly rely on quasi-legislative policy choices under conditions of scientific uncertainty, rather than pretending that their risk-assessment decisions are predicated on reliable scientific judgments, appellate courts might be less prone to accept arguments that agency analyses are irrational or flawed from a scientific perspective. There is no perfect way for administrators to protect their decisions against unsympathetic appellate review, but the current agency emphasis on "good science" invites judicial criticism of toxic risk assessments on grounds where the assessments are sure to be especially vulnerable.

IV. Integration of Science and Social Policy Judgments

EPA's carcinogen guidelines and an influential National Research Council study maintain that risk assessors should strive to make the best possible scientific judgments based on current knowledge and ordinarily should divorce these judgments from the economic, political, and ethical dimensions of regulation.[224] No one favors *bad* science, but this "good science" perspective is simplistic and potentially harmful in toxic contexts where the best available science is unreliable. When no consensus exists on how to resolve fundamental scientific uncertainties, policy considerations should and must influence agency choices on which provisional risk estimates to adopt. Explicit incorporation of social policy judgments into the risk-assessment process raises two related problems: Which types of policy criteria should be considered in risk-assessment analyses as well as risk-management decisions, and should the distinction between risk assessment and risk management be maintained once risk assessors employ social policy criteria to resolve scientific uncertainties.

A. *Applicable Social Policy Criteria*

Several types of policy criteria can guide risk assessments when an agency decides that the best available science is insufficient to yield reliable risk estimates. Some of these criteria may be evaluated once for each regulatory program and can provide the basis for generic treatments of recurring issues, while other material factors are linked to the particular characteristics of each toxic substance and must receive individualized treatments.

224. *See* EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,993 (1986); NATIONAL RESEARCH COUNCIL, *supra* note 1, at 151-52.

1. *Interpretation of Legislative Policies*

An agency choice among competing treatments of uncertainty on any scientific issue should be shaped by the policies in the applicable regulatory legislation. Legislatures have often recognized scientific uncertainties associated with toxic hazards and nonetheless required agencies to impose effective regulatory controls.[225] For example, the Occupational Safety and Health Act's legislative history observed "it is vital that when the Secretary sets an occupational health standard he do so on the basis of the best available evidence; it is not intended that the Secretary be paralyzed by debate surrounding diverse medical opinions."[226] In a similar manner, the DHS benzene assessment was conducted pursuant to a statutory mandate that provides "while absolute and undisputed scientific evidence may not be available to determine the exact nature and extent of risk from toxic air contaminants, it is necessary to take action to protect public health."[227] These legislative prescriptions do not offer a complete program for defining how agencies should resolve scientific uncertainties, but they do provide guidance that may help regulators develop their own systematic social policy responses. Explicit legislative mandates for protection against toxic substances despite the recognized presence of uncertainty should caution against agency adoption of "good science" requirements that in effect preclude control of most known or probable toxic hazards.

The Clean Air Act's treatment of hazardous air pollutants may provide another illustration of how regulators can shape risk-assessment practices in light of specific legislative policies. The Act requires national ambient air quality standards (NAAQS) to maintain an "adequate margin of safety," while standards for hazardous air pollutants must provide an "ample margin of safety to protect the public health."[228] The unequivocal language on hazardous air pollutants indicates that Congress intended to place a high, and possibly absolute, priority on assurance of public protection in this regulatory context.[229] The legislative history of the 1977 Amendments expressed congressional dissatisfaction with the failure of NAAQS to include safety margins equal to those in other environmental control programs, such as radiation standards.[230] The Committee com-

225. *See, e.g.*, Latin, *supra* note 5, at 381-83; Latin, *supra* note 12, at 603-05, 612, 620-21.

226. SENATE COMM. ON LABOR AND PUBLIC WELFARE, LEGISLATIVE HISTORY OF THE OCCUPATIONAL SAFETY AND HEALTH ACT of 1970, at 848 (1971) [hereinafter OSH ACT LEGISLATIVE HISTORY].

227. DHS BENZENE OVERVIEW, *supra* note 166, at 2 (quoting CAL. HEALTH & SAFETY CODE §39650(e) (West 1986)).

228. *See* 42 U.S.C. §7409(b)(1) (1982).

229. *See* D. CURRIE, *supra* note 75, at §§ 3.26, 3.28.

230. *See* CONGRESSIONAL RESEARCH SERV., 95TH CONG., 2D SESS., 4 A LEGISLATIVE HISTORY OF THE CLEAN AIR ACT AMENDMENTS OF 1977, at 2573-95, 2674-78 (Comm. Print 95-16, 1978).

ments justified "adequate" safety margins of one to two orders of magnitude in response to scientific uncertainty about the health effects of widespread air pollutants.[231] It seems reasonable to conclude that hazardous air pollutant standards, which are at least equally subject to scientific uncertainty and are supposed to include "ample" safety margins, should be even more biased in the direction of protection than the NAAQS limits. This interpretation is consistent with Congress's special concern for protection against toxic hazards, as expressed in an array of regulatory statutes enacted during the past two decades.[232] Yet, EPA has not attempted to resolve the many uncertainties presented by regulation of toxic air or water pollutants systematically in light of this congressional preference.

Juxtaposition of the statutory meanings of "adequate" and "ample" safety margins suggests that conservative treatments of toxic risks would generally be more compatible with congressional intent than "mid-range" positions, and the same conclusion would apply to "best-current-guess" estimates unless they can be made with a high degree of scientific confidence. This interpretation does not mandate the most protective possible risk-assessment treatment on every contested issue, nor does it imply that the statutory language and legislative history are so clear that they preclude any degree of administrative discretion. The statutory preference for "ample" safety margins is, however, sufficiently palpable to require that EPA provide a persuasive rationale, not merely conclusory assertions, in hazardous pollutant contexts where the Agency selects less conservative positions. Although EPA's carcinogen guidelines adopt protective treatments on some important issues, they do not incorporate a consistently conservative bias[233] nor provide any explanation of how "ample" safety margins will be assured for carcinogenic air pollutants. Indeed, the guidelines ignore this statutory provision and instead require individualized analyses based on the best available, albeit not necessarily "good," scientific evidence and theories.

EPA's only reference in its benzene proceedings to an "ample margin of safety" appears to distort the statutory meaning of the clause. After deriving unit-risk and exposure estimates, the Agency identified the degree of pollution reduction available through application of the "best available" control technology (BAT) for coke by-product recovery plants.[234] EPA compared the protection achievable by its designated BAT with the incre-

231. *See id.*; D. CURRIE, *supra* note 75, at § 7.11.
232. *See* Latin, *supra* note 5, at 392-93 & n.449.
233. *See supra* text accompanying notes 44-68.
234. *See* EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,533-35 (1984).

mental safety offered by more stringent measures[235] and predicted that better-than-BAT control technologies would reduce the risk only from 0.19 to 0.06 leukemia cases per year at considerably higher cost.[236] The Agency then concluded:

> Because of the relatively small health benefits to be gained with the additional costs and the potential adverse economic impacts on some firms of requiring the [better than BAT] option, EPA considers the risks remaining after application of BAT not to be unreasonable. For this reason, EPA judged the level of control selected as BAT to provide *an ample margin of safety* and decided not to require a more stringent level of control than BAT for coke by-product recovery plants.[237]

EPA's risk assessments for BAT and for more protective control technologies used identical treatments of the unit-risk factor, pollutant dispersion patterns, and exposed populations; the only difference was in projected discharge levels. The Agency did not include an ample margin or any systematic safety margin in its scientific assessment of benzene risks and cited the statutory clause *only* in its risk-management analysis of regulatory cost-effectiveness. Even if EPA may consider costs in setting hazardous air pollutant standards,[238] the risk-management analysis should be

235. The Agency explained its analytical process as follows:

> Best available technology for new and existing sources is technology which, in the judgment of the Administrator, is the most advanced level of control considering the economic, energy, and environmental impacts and any technological problems associated with retrofitting of existing sources.
>
> After selecting BAT, EPA identified a level of control more stringent than BAT and evaluated the incremental reductions in health risks obtainable against the incremental costs and economic impacts estimated to result from the application of the more stringent control level. This provides a comparison of the costs and economic impacts of control with the benefits of further risk reduction.

Id. at 23,533. This treatment is confusing and appears incompatible with the legislative intent. If a more stringent control technology is technologically and economically feasible, *see id.* at 23,535–37, how did EPA designate less stringent technology as the most advanced and best available BAT level? It appears to me that the Agency first used cost-effectiveness criteria to identify a control level somewhat arbitrarily as BAT, and then used a further cost-effectiveness comparison to justify its decision not to impose better control measures. I believe Congress imposed either an absolute priority on safety from toxic substances or a feasibility test, not a cost-effectiveness balancing process, for most toxic control programs. *See, e.g.*, American Textile Mfrs. Inst. v. Donovan, 452 U.S. 490, 508–12 (1981); NATIONAL RESEARCH COUNCIL, *supra* note 1, at 42–48; Latin, *supra* note 5, at 320.

236. The Agency estimated that requiring the more stringent set of controls would entail capital costs of $131 million and marginal operating costs of $37.5 million per year. *See* EPA Proposed Standards for Benzene Emissions, 49 Fed. Reg. 23,522, 23,537 (1984).

237. *Id.* (emphasis added).

238. Although EPA may not consider regulatory costs in setting ambient pollution limits under the NAAQS, the Agency construed the hazardous air pollutants provision to allow balancing of control costs and regulatory benefits. *See* Goldstein, *supra* note 103, at 298. This administrative interpretation implausibly presumes that Congress was more concerned about the economic impacts of regulatory standards for especially dangerous toxic discharges than it was with regard to the costs of less

performed after risk estimates are adjusted to reflect ample safety margins under conditions of scientific uncertainty. Although agencies must be given latitude to interpret their statutory mandates, I do not see how EPA could plausibly contend that its selective treatment of safety margins in the benzene proceedings was consistent with the applicable congressional balance of interests and priorities in the organic statute.

2. *Cost-Effectiveness of Individualized Analyses*

If the effectiveness of toxic substances control programs is considered on a synoptic level, regulatory agencies must acknowledge that their achievements have fallen far short of legislative intentions. Indeed, these programs often suffer from bureaucratic paralysis and are invariably more expensive and time-consuming than Congress or the agencies themselves expected. Notwithstanding its criticisms of the administrative process, the National Research Council rightly concluded that "the basic problem in risk assessment is the sparseness and uncertainty of the scientific knowledge of the health hazards addressed, and this problem has no ready solution."[239] Yet, agencies have seldom examined the effectiveness of their risk-assessment procedures in light of this fundamental problem. Given the inherent complexity of toxic hazards and severe constraints on agency resources, regulators must consider which analytical procedures are cost-effective and which scientific issues are worth assessing repeatedly in individualized proceedings.

EPA's carcinogen guidelines are almost entirely lacking in this form of self-analysis. The guidelines provide no indication of which risk-assessment issues are especially difficult or expensive to address. They do not specify which analytical issues cannot now be resolved in a reasonably reliable manner due to the absence of any scientific consensus and which issues have been raised repetitively but inconclusively in prior regulatory proceedings. They do not identify which risk-assessment issues and procedures are likely to enable obstructive behavior by regulated parties. For example, EPA and DHS were required in their 1984 proceedings on benzene risks to assess debatable extrapolation theories and negative epidemiological data that had been examined in detail and rejected by OSHA, CAG, and IARC in earlier analyses. Given the great difficulty in regulating any toxic substance, it is doubtful that agencies should assess in each instance whether safe threshold levels exist, whether one speculative

hazardous NAAQS pollutants. Even if EPA is correct that economic effects should be considered, that objective can be achieved under the feasibility test commonly used for toxic controls and need not justify the Agency's adoption of a cost-effectiveness test based on speculative risk estimates.

239. NATIONAL RESEARCH COUNCIL, *supra* note 1, at 6.

extrapolation theory should be preferred over competing models, and whether benign tumors should be excluded from findings of animal studies. No doubt these and many other recurring issues are relevant to attainment of the best possible science, but individualized assessments of all material scientific issues in all toxic control proceedings may preclude achievement of adequate and timely protection for exposed populations.

OSHA's generic cancer policy under the Carter Administration attempted to foreclose discussion of many contested issues that the Agency had previously addressed.[240] EPA's current carcinogen guidelines go to the opposite extreme by mandating individualized analyses of any potentially relevant scientific theories or data in each proceeding. Regulators might choose an intermediate position in which generic presumptions against certain kinds of theories or evidence could be rebutted by a credible showing that a scientific consensus has emerged on a previously contested issue. Risk assessors should recognize that their treatments of recurring issues and uncertainties have important implications for the scope, cost, and timing of toxic substances regulation. It is unclear whether EPA's disregard of such factors in the carcinogen guidelines represents an instance of scientific tunnel vision or a deliberate attempt to impede effective regulation, but the guidelines and accompanying explanatory statements never question the utility of the Agency's "good science" focus.

3. *Potential for Catastrophic Miscalculations*

Regulators could increase the conservative bias in their risk estimates when a particular toxic substance may have catastrophic effects if it proves more potent than the agency assessment anticipates. The presence of several individualized circumstances might support this form of social policy judgment:

a. *Widespread Population Exposures*

Some chemical usage and dispersion patterns entail significant exposures for only a relatively small number of workers or residents near pollution sources, while other hazardous substances may endanger millions of people. If the "best-current-guess" prediction underestimates actual risks by more than two orders of magnitude for pollutants where exposure is limited or localized, the result may be "only" a few dozen unexpected fatalities. In contrast, similar mistakes in estimation of the risks from widespread toxic exposures may have catastrophic effects, such as those

240. *See* OSHA Generic Cancer Policy, 45 Fed. Reg. 5002 (1980); Latin, *supra* note 5, at 364-66.

associated with asbestos and DES. For example, because benzene is a ubiquitous chemical, Dr. Albert noted that a risk-assessment error of several orders of magnitude "would be a national disaster."[241] Most regulatory assessments now treat a predicted high risk for a small group as statistically equivalent to a predicted small risk for a large population, but that approach does not adequately consider the possible consequences of mistakes in the agency's determinations.

It is widely recognized that susceptibilities of individuals and population subgroups to toxic hazards vary widely, but there is currently no accepted methodology for tailoring risk estimates in response to those differences.[242] Present risk estimates are usually based on dose-response data derived from epidemiological studies of the entire population or of white male workers. Agency decisionmakers may, however, choose to adjust risk estimates in order to provide additional protection for unusually vulnerable subgroups, as in the cases of exposure of children to high lead concentrations or of pregnant working women to certain hazardous industrial chemicals. Again, this precautionary judgment would reflect social policy considerations in light of the possible consequences of agency mistakes under conditions of scientific uncertainty, rather than the current practice of treating risk assessment purely as a function of "good science."

b. *Absence of a Long Historical Record of Exposures*

Some toxic substances, such as benzene, have been in common use for decades at higher exposure levels than are now prevalent. This historical record reduces the chance of catastrophic risk-assessment errors because hazards of epidemic proportions presumably would already have manifested themselves. In contrast, many substances are introduced each year that may eventually have toxic effects, and long latency periods may conceal those hazards for decades. Regulators might increase the conservative bias in their risk estimates for substances that lack a long historical record of exposures and related health effects. Yet, agencies seldom consider this factor in their scientific risk-assessment deliberations. The legislative history of the Occupational Safety and Health Act, for example, emphasized the uncertainty created by "literally hundreds of new chemicals" that are "introduced into industry at a much faster rate than the present meager [regulatory] resources . . . can keep up with."[243] OSHA's generic cancer

241. *Comments*, *supra* note 46, at 329.

242. *See* EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986); EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,479, 23,485 (1984); DHS BENZENE REPORT, *supra* note 39, at 52-53; Comments in 19 BANBURY REPORT, RISK QUANTITATION AND REGULATORY POLICY, 337, 340 (D. Hoel, R. Merrill & F. Perera eds. 1985) (comment of Dr. Perera).

243. OSH ACT LEGISLATIVE HISTORY, *supra* note 226, at 142, 849; *see also id.* at 159-60, 415, 517, 1048-49.

policy and individual risk assessments, however, have not made distinctions in the treatments of longstanding and new chemicals.[244]

c. *Evidence of Unusual Potency*

Regulators may occasionally receive evidence that a substance is unusually hazardous prior to their acquisition of sufficient data for a reliable dose-response assessment. In some instances, the substance under investigation may have a close chemical resemblance to another substance whose hazardous effects have been clearly documented, or it might yield positive results in short-term *in vitro* tests. In other cases, data from animal studies may reveal an especially high degree of toxic potency. Risk assessors could decide on the basis of these preliminary but suggestive indications of severe toxic hazards to increase the conservative bias in risk estimates derived from incomplete data. For example, CPSC tried to regulate formaldehyde after one animal study found that the substance may be an unusually potent carcinogen, while the Fifth Circuit's contrary decision left millions of people exposed to a hazard of unknown but potentially serious dimensions.[245]

4. *Ability to Resolve Uncertainty*

In practice, agencies seldom commence regulatory proceedings until considerable evidence has accumulated that a substance may be hazardous. When sufficient information or public controversy exists to justify an expensive risk-assessment hearing, agency experts usually consider whatever data happen to be available at the time. As Dr. Albert noted "[r]isk assessment is a passive operation which accepts the data that comes in."[246] In some instances, however, agencies may be able to identify ongoing scientific studies or to sponsor collection of data on acute toxic effects, prevailing exposure patterns, or other material issues. In such cases, regulators might adopt interim strategies on the assumption that specific uncertainties can be resolved in the near future. They might, for example, allow a substance under investigation, such as a newly developed drug, to be used when no substitute is available but not if its primary advantage is lower costs. Because risk assessors typically cannot predict the outcome of

244. Some regulatory programs, such as those for drugs and pesticides, require industry to test substances for possible toxic effects before they can be marketed. Although these regulatory requirements nominally place the burden of proof on the new chemical, there is little indication that EPA and other responsible agencies vary their scientific risk assessment practices in a systematic manner based on the distinction between familiar and new substances.

245. *See supra* text accompanying notes 203-09.

246. Comments, *supra* note 46, at 325.

scientific research with assurance, this type of hedging strategy clearly entails a problematical policy choice to accept some risks on a tentative basis in return for the social benefits associated with use of the toxic substance.[247] This criterion could, however, facilitate abuse of discretion by agency officials because it may allow amorphous trade-offs influenced by political or economic pressures.[248] Regulators should therefore be required to explain in detail which studies are likely to resolve particular scientific uncertainties, and why the anticipated findings may warrant interim regulatory strategies.

It is equally important for agencies to explain why many scientific uncertainties cannot be eliminated in the immediate future.[249] Given the propensity of affected parties to challenge agency determinations on every possible ground, regulators should preemptively acknowledge their inability to provide dispositive scientific answers to many important questions. It is true that this degree of candor may sometimes lead to adverse appellate decisions, but experience indicates that judicial review of past agency practices has been at least as intrusive.[250] Only through explicit consideration of prevailing uncertainties and the policies appropriate to determine their social consequences can regulators hope to persuade legislatures and appellate courts that requirements for precise quantitative assessments of toxic risks are unrealistic and invite either agency paralysis or unreliable speculation. The social policy criteria identified in this Part can justify adoption of more (or less) protective risk estimates despite continuing uncertainty about toxic risks, but regulators must emphasize the policy-oriented bases of their decisions and must explain why they frequently cannot provide reliable scientific conclusions without indefinitely delaying the imposition of protective controls or compromising other statutory objectives.

The National Research Council study of risk-assessment problems advised regulatory agencies to adopt generic approaches for risk-assessment problems on the grounds that uniform science policy guidelines

247. For a discussion of whether this type of social benefits criterion should be reserved for risk-management analyses or should be incorporated in risk assessments, see *infra* text accompanying notes 252-63.

248. Moreover, the history of environmental control programs indicates that interim pollution control standards often remain in effect indefinitely as a result of agency inertia and higher regulatory priorities. *See* Latin, *supra* note 11, at 1305-07, 1319-20.

249. Although scientists have acquired useful information about diverse carcinogenesis processes during the past decade, few if any of the critical uncertainties have been resolved. *See, e.g.*, NATIONAL RESEARCH COUNCIL, *supra* note 1, at 28-37; EPA DECISION MAKING, *supra* note 5, at 65; Preuss & White, *supra* note 43, at 335. Many of these uncertainties stem from inadequate scientific understanding rather than inadequate data, and cannot necessarily be remedied by well-funded research programs. *See* Latin, *supra* note 11, at 1283.

250. *See supra* text accompanying notes 203-23.

"could help separate risk assessment from risk management considerations, improve public understanding of the process, foster consistency, and prevent oversights and judgments that are inconsistent with current scientific thought."[251] It is often advisable for regulatory agencies to rely on generic treatments of recurring scientific issues, but generic policies cannot resolve all scientific and social policy questions in each toxic risk-assessment proceeding. Particularized circumstances, such as those pertaining to the potential for catastrophic errors or the likelihood that specific uncertainties can be resolved, would preclude resolution of all uncertainties in a consistent fashion. Requiring agencies to provide cogent reasons for their treatments of varied types of scientific uncertainty may be more realistic and more important than a high degree of uniformity in risk-assessment outcomes.

B. *Interaction of Risk Assessment and Risk Management*

Most discussions of risk assessment stress the need for scientists or regulators to identify significant uncertainties and to explain the assumptions used to resolve them.[252] EPA's carcinogen guidelines acknowledge "in every quantitative risk estimation that the results are uncertain"[253] and then provide:

> Whichever method of presentation is chosen, it is critical that the numerical estimates not be allowed to stand alone, separated from the various assumptions and uncertainties upon which they are based. The risk characterization should contain a discussion and interpretation of the numerical estimates that affords the risk manager some insight into the degree to which the quantitative estimates are likely to reflect the true magnitude of human risk, which generally cannot be known with the degree of quantitative accuracy reflected in the numerical estimates . . . . Major assumptions, scientific judgments, and, to the extent possible, estimates of the uncertainties embodied in the assessment are presented.[254]

This aspirational proviso does not ensure that risk managers can make *meaningful* judgments on the basis of an agency's recitation of unresolved issues. An extensive list of scientific uncertainties would not, by itself, be very useful for agency risk managers or for public understanding of toxic hazards. Neither is the common agency practice of pairing discussions of

251. NATIONAL RESEARCH COUNCIL, *supra* note 1, at 162; *see also id.* at 69–82.
252. *See id.* at 153–54, 164, 169.
253. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,997 (1986).
254. *Id.* at 33,999.

uncertain factors that may induce overestimation of risks with other indeterminate factors that may understate risks. EPA's carcinogen guidelines and benzene regulations fail to explain adequately the possible dimensions and ramifications of many uncertainties, and they clearly do not provide enough information for risk managers to determine the appropriate regulatory consequences of scientific uncertainty. This observation is not intended as a criticism of past agency practices because the proposed allocation of responsibilities is fundamentally unrealistic. Only risk assessors are likely to understand the dimensions and implications of particular uncertainties sufficiently well to make effective social policy judgments on how those uncertainties should be resolved. If risk assessors cannot accomplish this task, risk managers could seldom be expected to make informed choices on the same issues.

In order to communicate a sense of the uncertainties in risk-assessment predictions, for example, some agency risk assessors believe their findings should be expressed as ranges of potential hazards at varying exposure levels, while other experts prefer discrete point estimates.[255] EPA's 1984 benzene regulation observed:

> EPA has concluded that the presentation of the risk estimates as ranges does not offer significant advantages . . . [R]anges for benzene make risk comparisons among source categories more difficult and tend to create a false impression that the bounds of the risks are known with certainty. For these reasons, the benzene risks . . . are presented as point estimates of the leukemia risk.[256]

EPA's carcinogen guidelines, in contrast, provide that a range of upper- and lower-limit risks should be explicitly stated in the assessment whenever possible.[257] The guidelines do not explain why EPA reversed its previous position on risk ranges and never address the problems identified in the quoted passage. The guidelines also concede that an "established procedure does not yet exist for making 'most likely' or 'best' estimates of risk within the range of uncertainty defined by the upper- and lower-limit estimates."[258] In view of the many uncertainties in toxic risk assessment, it is unclear whether risk ranges or point estimates create *more* of "a false impression that the bounds of the risks are known with certainty." It is doubtful that risk managers can select risk estimates intelligently from

255. For example, The DHS report derived a range of possible benzene risks that varied by a factor of about eight. *See supra* text accompanying notes 171–173.

256. EPA Regulation of Benzene, 49 Fed. Reg. 23,478, 23,493 (1984).

257. *See* EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,998 (1986).

258. *Id.* The caveat in the text invites skepticism about how accurate the upper- and lower-limit boundary estimates are likely to be.

within ranges in which the upper and lower bounds may differ by several orders of magnitude, and yet the apparent precision of specific point estimates is likely to be misleading.

There is little if any evidence that agency administrators carefully tailor their decisions in light of the uncertainties identified in the risk assessment. Instead, risk managers generally accept risk estimates as gospel and then examine economic and political issues relevant to the toxic substance at issue. Indeed, the scientists most responsible for conducting EPA risk assessments evidently recognize that risk managers accept their quantitative estimates uncritically regardless of accompanying qualifiers which describe unresolved issues. Consider the following exchange at a 1985 symposium on risk assessment:

> Professor Merrill: Dr. Albert, you indicated the program officers at EPA were insisting on a mathematical expression of risk. Were they insisting on the mathematical integration of those qualifiers?
>
> Dr. Albert: They didn't really know. It is evident that if there is weak qualitative evidence for carcinogenicity, it ought to be factored into the quantitative assessment. If one regards the public health impact of cancer in terms of the excess estimated numbers of cancer cases, then that estimated impact should be clearly reduced by weak qualitative evidence.
>
> Dr. Preuss: If you could present them separately and discuss them and give the person making the decision some idea of the strengths and the weaknesses on both sides.
>
> Dr. Albert: All they seem to want to know is "Is it a carcinogen or isn't it a carcinogen?"
>
> Dr. Preuss: And you're going to tell them it's 0.03 of a carcinogen?
>
> Dr. Albert: No. The bottom line here is the estimate of the number of cancer cases, and that is reduced in proportion to some arbitrary scaling of the probability that the agent really is a human carcinogen.[259]

In other words, Dr. Albert's Carcinogen Assessment Group may reduce its best-guess risk estimate by some "arbitrary scaling" to reflect the presence of scientific uncertainty, but he does not expect the Agency's risk managers to understand that adjustment. Under this "good science" procedure, EPA risk assessors do not explicitly tailor risk estimates in light of social policy judgments and risk managers do not modify the estimates at all. Dr. Preuss observed to Dr. Albert later in the symposium discussion

259. Comments, *supra* note 46, at 328–29.

"[m]y concern, Roy, is that all of the qualifications get lost and we focus on that number."[260] These candid statements by Agency experts indicate that social policy criteria must be incorporated into the risk-assessment process if they are to have any significant effect on the choice among competing estimates under conditions of scientific uncertainty. As previously emphasized, alternative predictions may vary by several orders of magnitude, and a risk assessor's selection of a speculative estimate can determine the ultimate regulatory decision irrespective of any social policies and values adopted in the risk-management process.

For example, analysis of whether adequate substitutes exist for a potentially toxic product or process is usually regarded as a risk-management function. The social value of a toxic chemical is determined in part by the availability of safer alternatives, and the existence of reasonable product substitutes would clearly be an important element in the risk-utility balancing comparison that often forms the heart of risk-management deliberations. If the risk manager is provided with a prediction that the substance under investigation poses only a minuscule hazard, the administrator would be unlikely to regulate that substance although safer substitutes are in common usage. Yet, the agency's risk estimate may be unreliable from a scientific perspective, even if it is the best-current-guess, and adoption of different risk-assessment assumptions could suggest a much greater danger. Because risk managers are seldom equipped or disposed to modify risk estimates, it may be appropriate for risk assessors to increase the conservative bias in their estimates when available substitutes could achieve reasonably equivalent functions.[261] Thus, both risk assessors and risk managers should evaluate social consequences and policy criteria in order to perform their respective functions. This conclusion conflicts with the position in EPA's carcinogen guidelines that risk estimates should be derived "independently from considerations of the consequences of regulatory action."[262]

This Article's contention that agencies cannot wait until the risk-management stage of toxic regulatory proceedings to address the social ramifications of scientific uncertainty does not suggest that the distinction between risk assessment and risk management should be abandoned. In

260. Comments, *supra* note 242, at 340.

261. For example, after the Fifth Circuit's decision on urea-formaldehyde foam insulation millions of people were subjected to low-level formaldehyde exposures despite the availability of other forms of nontoxic insulation. When a toxic substance's primary benefit is a marginal cost advantage over substitute products, agencies should not necessarily adopt the same risk-assessment treatments that they employ in the context of new drugs, pesticides, or other hazardous materials that offer distinctive benefits. Yet, most risk-assessment treatments do not consider this type of distinction because it is based on social policy, not scientific, considerations.

262. EPA Carcinogen Guidelines, 51 Fed. Reg. 33,992, 33,993 (1986).

some toxic contexts, risk assessors can provide reliable estimates based on generally accepted scientific principles.[263] There is no reason why risk assessments in such contexts should be modified in response to policy criteria when "good science" judgments can really be grounded on valid science. Moreover, risk managers must address economic, political, and ethical factors relevant to each toxic substance even if the social policies applied in the risk-assessment stage are incorporated in generic treatments of recurring scientific issues. To the extent the conventional distinction represents a real rather than symbolic division of decisionmaking responsibilities, risk managers should retain the ultimate authority to determine the scope of toxic regulations. Nevertheless, risk assessment inescapably plays a central role in the toxic substances regulatory process and is too uncertain to be treated exclusively as an exercise in "good science."

Conclusion

Environmentalists attack the risk-assessment process because they believe it frequently produces unreliable estimates of toxic hazards and because it is subject to manipulation by industrial dischargers and government bureaucrats. However sympathetic one may be to these objections, which surely have ample basis in past regulatory experience, society cannot feasibly eliminate all carcinogenic risks nor enjoin use of all toxic substances. Society must therefore develop some rational method for deciding which risks are unacceptable and for allocating scarce regulatory resources. Notwithstanding the risk-assessment uncertainties and analytical shortcomings emphasized in this Article, it is unlikely that regulators should or could eliminate attempts to estimate the dimensions of diverse toxic hazards. Moreover, after a decade of intrusive appellate decisions and political emphasis on cost-effectiveness justifications, risk-assessment procedures are firmly embedded in the federal regulatory agencies responsible for toxic substances control. In short, we are largely past the question of whether to assess toxic risks, and must now determine how to improve the quality of regulatory decisionmaking in order to promote the legislative goals and social values reflected in environmental protection programs.

There have been any number of discussions in the legal and public policy literature about mechanisms intended to improve the efficiency of environmental regulation through reliance on cost-benefit, risk-utility, or

263. In the case of some noncarcinogenic toxic substances, such as cotton dust and lead, that have been in use for many decades and produce chronic effects after long-term exposures, scientists may be able to obtain reliable epidemiologic data and identify reasonably accurate dose-response relationships. *See* American Textile Mfrs. Inst. v. Donovan, 452 U.S. 490, 505 n.25 (1981).

cost-effectiveness analyses. These strategies are fundamentally dependent on the quality of regulatory risk assessments, and yet most commentators treat the risk-assessment process as a "black box" and assume that agencies must always produce the best available scientific predictions regardless of the imperfect state of the art. Regulators cannot avoid determining the social and legal consequences of scientific uncertainty, but there is no assurance that they will do so in a systematic manner responsive to legislative objectives. Not only have agencies generally failed to explain the implicit policies that shape their risk-assessment practices, but it also appears that regulators often have not recognized the social ramifications of their own treatments. Moreover, current risk-assessment practices are less consistent and less reliable than agency scientists typically concede.

Because predictions of toxic effects generally cannot be grounded on reliable scientific judgments, social policy criteria must play an influential role in the choice among competing risk estimates. Once we recognize that toxic substances regulation requires a panoply of policy determinations to supplement provisional scientific judgments, it is essential that risk-assessment agencies explicitly consider the social ramifications of scientific uncertainty, strive for analytical coherence in their treatments of currently indeterminate issues, and clearly explain the principles, practices, and values underlying particular estimates of toxic hazards.