THE NEW PERSPECTIVE ON HEALTH: PREVENTION AND HEALTH STANDARDS

Giandomenico Majone

March 1981
PP-81-6

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

This paper is an expanded version of the Kaiser Lecture given at Yale University, on October 22, 1980. It will appear in the volume The New Perspective on Health edited by Theodore R. Marmor and to be published in 1981.

An appendix has been added. In it, some questions raised in the paper, relating to the determination of the dose-response function and the use of cost-benefit analysis in standard-setting, are discussed at a more technical level.

THE NEW PERSPECTIVE ON HEALTH: PREVENTION AND HEALTH STANDARDS

Giandomenico Majone[*]

1. Standard-setting in a comparative perspective

The growing debate over preventive approaches to health problems cannot proceed far without encountering issues related to environmental and occupational health. A country's commitment to prevention may be judged from the way it goes about protecting the living and working environment of its citizens. In this area of public policy significant changes have taken place, during the last decade, in all major industrialized countries. Nowhere these changes have been more remarkable than in the United States. Here a series of legislative enactments--the Occupational Safety and Health Act of 1970, the Amendments to the Clean Air Act passed in the same year, the Safe Drinking Water Act of 1974, and the Toxic Substances Control Act of 1976, to mention only some--have brought about major shifts from previous policies based on decentralized control and voluntary compliance toward compulsory regulation set at the national level.

In all these legislative enactments, standards appear as the most important policy tool for the prevention of accidents, ill health, and environmental degradation. In

---

[*]Faculty of Economics and Social Sciences, University of Calabria, Italy, and International Institute for Applied Systems Analysis, Laxenburg, Austria.

The present essay is an expanded version of the Kaiser Lecture given at Yale University, October 22, 1980.

fact, American regulatory philosophy, especially in the field of occupational health and safety, is moving away from the approach prevailing in West Europe--where standards are typically used as general guidelines to be flexibly interpreted by inspectors in each particular case--, and comes closer, in some respects, to the practice of the Soviet Union, where standards are embodied in legally binding codes, and public health is defined as the science of setting "optimal" health standards.

Standards play different roles in different countries. Also their numerical values vary a good deal from country to country. For example, it is well known that Soviet environmental and occupational health standards are typically much more stringent than those used in the United States or West Europe. Such cross-national variations raise a number of questions about the nature of the standard-setting process, as well as the cognitive, philosophic, and institutional factors that shape it.

In this as in other areas of policy making, comparative analysis represents a convenient method for exposing hidden assumptions and unquestioned beliefs. More important, comparative analysis shows that the setting of health standards, far from being an almost mechanical process that can be safely delegated to technicians, in reality represents a microcosm in which national traditions, philosophies, attitudes, and institutions are faithfully reflected.

The insights contributed by a comparative perspective have practical implications as well as intellectual interest.

Consider, for example, the fact that in the United States, with its enormous scientific, technical, and financial resources, no more than 500 chemicals can be tested each year because of the limited availability of trained toxicologists, laboratory facilities, and test animals. This is barely sufficient to keep up with the flow of new chemicals, let alone investigating the existing stock of well over 50,000 chemicals already in commercial use. International cooperation in toxicological testing would have obvious benefits but serious, if ill-understood, differences in methodology, risk philosophies, and regulatory approaches make cooperation difficult, and even reduce the value of the limited amount of information that is available.

Again, all industrialized and most developing countries make extensive use of environmental and health standards, but much of the research from which these standards are derived is done in a handful of countries--primarily the United States and the Soviet Union. However, independent verification and intelligent adaptation of the research results to particular national situations are difficult because of the abundance of implicit assumptions and the lack of standardized procedures. The result is a mechanical adoption of "foreign" standards, barely disguised by ad hoc manipulations of safety factors and other rules of thumb. Errors of fact and logic in the original derivations are propagated and magnified in the process .

In the field of environmental and occupational health we may have reached a point were fine-tuning of intrinsically

inadequate regulatory mechanisms (such as requiring formal cost-benefit analyses in setting compulsory standards) can only serve to impede truly innovative thinking. Probably the most important contribution that comparative analysis can make is disclosing the variety of institutional solutions that are possible, and have been used or at least proposed in different national contexts.

In the following pages I shall explore some sources of variation in the way health standards are derived and used in different countries--differences in cognitive paradigms, in regulatory philosophies, and in enforcement procedures. This analysis will provide the empirical support for some comments about the uses and limitations of health standards. The possibility of voluntary standards and self-regulation will be discussed in the latter part of the paper.

2. The conceptual basis of standard-setting

The derivation of health standards in various countries reflects, first of all, differences in the definition of what is a state of health, as well as conflicting views concerning the degree to which the defense mechanisms of the body can be safely drawn upon to offset insults from toxic agents and pollutants.

Toxicological procedures used in the West rely on the idea that no threat to health exists so long as the exposure does not induce a disturbance such as to overload the normal protective mechanisms of the body. For example, the United

States National Academy of Sciences defines non-adverse health effects as

> (1) changes that do not result in impairment of functional capacity or the ability to compensate for additional stress;
>
> (2) changes that are reversible following cessation of exposure if such changes occur without detectable decrements in the ability of the organism to maintain homeostasis; and
>
> (3) changes that do not enhance the susceptibility of the organism to the deleterious effects of other environmental influences.[1]

According to Soviet biological philosophy, any change in response to stimulus represents an unacceptable deviation from normal conditions, and any concentration, however small, places an undesirable toxic or nuisance stress on the organism. Thus, a potential for ill health is said to exist as soon as the organism undergoes the first detectable change of whatever kind from its normal state.[2]

To better visualize these conceptual differences, imagine the familiar dose-response curve as being subdivided into three zones: an upper zone corresponding to high doses of a toxic substance, where ill effects due to exposure are clearly detectable; a compensatory zone where the body adjusts to the stresses imposed by lower levels of exposure, but at some cost; finally, a lower, homeostatic zone where the adjustments are automatic. According to the biological philosophy prevailing

in the West, the defense mechanisms in the compensatory zone, as well as the normal adaptive processes in the homeostatic zone, can be safely drawn upon, within limits, to offset the levels of stress imposed by minimum exposure to hazardous agents at the workplace and in the environment, just as they are called upon to counter the wear and tear of ordinary life. Hence, dose-response relationships are extrapolated downward from the zone of demonstrable health burdens to a point of "non-detectable" ill effects.(3)

Soviet toxicologists start at the other hand of the dose-response curve, moving upward from zero dose and a corresponding initial benchmark of normality in the test organism. The permissible level of exposure is established below the lowest dose needed to induce a statistically significant difference from the normal state, as revealed by highly sensitive measures of behavioral response. The assumption underlying this procedure is that the protective mechanisms in both homeostatic and compensatory zone should be kept in reserve to take care of unexpected insults, and their effectiveness should not be weakened by the continuous demands of stress knowingly permitted in the environment or at the workplace.

Neither the Soviet nor the western position can be dismissed as being unreasonable or contrary to known biological laws, but the practical implications in terms of acceptable levels of exposure are vastly different in the two cases. The official goal of Soviet standard-setters is a zero level of exposure. By contrast, goals of zero exposure have not been seriously discussed in the United States or in other countries

in the West, except for radiation protection and for carcinogens in the workplace.[4]

Conceptual differences concerning the nature of health and the adaptive capacity of the human organism, are magnified by differences in research techniques. Soviet toxicologists place major emphasis on the study of the effects of toxic agents on the nervous system. Central nervous system sensitivity (conditioned reflexes, electroencephalogram) and reflex responses (changes in heart and respiratory rate, in blood pressure, and so on) play a central role in standard setting. In the words of a Soviet expert,

> We attach great significance to chronic changes in the higher nervous activity of animals under the influence of toxic substances in the air they breathe. We believe that changes in the functioning of the cortex of the cerebral emespheres occur very early, even with small concentrations, since the cerebral cortex is highly sensitive to the effects of external factors in the environment... One of the early manifestations of the influence of various chemical substances on the higher nervous system is the development of phasic states. Later, disinhibition of differentiation occurs, then individual reflexes begin to disappear and finally none of the reflex pattern is left. When the animal is more severely affected, the natural conditioned reflex to sight and smell of food disappears.[5]

Because of this preoccupation with the role of the higher nervous system as controller of all bodily activity, considerable more importance is given to the pathology of this system than is the case for Western studies.[6] Incidentally, the interest of Soviet toxicologists for nervous system testing and reflex behavior can be explained by the enormous influence of Pavlovian theories on all domains of Soviet medicine.[7] In particular, the insistence on nervous system testing is justified by reference to Pavlov's theory that living organisms

adapt to their environment by means of two nervous mechanisms: the unconditioned reflexes for the permanent features of the environment, and the conditioned reflexes for the temporary (conditional) features.

American and European scientists, while not fully convinced that nervous system testing necessarily provides more sensitive indicators of toxic action, agree that sophisticated measurements of nervous system effects should be a more important part of toxicological testing in the West.(8)

Another interesting methodological difference is the limited role which epidemiology seems to play in standard setting in the Soviet Union. In the West, and particularly in the United States, epidemiology has historically provided important, and sometime decisive, evidence on which standards have been based; although there are indications that its role may be decreasing relative to toxicological testing. For the Soviets, on the other hand, epidemiological studies represent a form of human experimentation in which prior toxicological tests and subsequent prevention have failed. In short, epidemiological studies represent a reactive rather than a preventive approach. Moreover, epidemiological studies abroad, showing health effects at higher concentrations than those allowed in the Soviet Union, encourage continued faith in the value of the traditional approach, based on toxicological evidence largely derived from nervous system testing, that has led to the lower Soviet standards.(9)

3. Differences in regulatory philosophies

Health standards, Soviet authorities maintain, should be based on health effects alone, without regard to the availability of adequate control technology, to economic feasibility, or even to the ability to adequately measure the concentrations in practice. A currently unattainable standard can still represent a guideline for enforcement and an incentive for future research in control technology. Conversely, since technically or economically attainable concentrations will coincide with harmless concentrations only by chance, standards based on considerations of economic or technical feasibility "can act only as an obstacle to the search for better techniques, ... they sanction what has already been achieved without stimulating new technical advances."(10)

Since "scientifically based" standards cannot always be achieved, the Soviets use also secondary ("sanitary") standards that may modify, for a limited period of time, the requirements set by the primary ("hygienic") standards. Professor V. A. Rjazanov, a leading toxicologist, distinguishes the two types of standards (in the context of air pollution) in the following terms:

> Hygienic standards ... must in themselves reflect the scientifically based ideal towards which we must strive in order to ensure that the public is not subject to unfavorable effects from air pollution. This ideal cannot be achieved always and everywhere at a given time. Therefore, alongside the general hygienic standards for maximum permissible concentrations, there may be sanitary standards of a temporary character, serving the needs of the moment. They may modify for a defined period the requirements for cleanliness of the external atmosphere, taking into account economic and technological factors... Such air pollution

> standards are permissible temporarily, but should be abandoned after a certain period, during which the condition of the air must be brought into conformity with the hygienic standards. If this approach is adopted, hygienic standards for the cleanness of the external air will not be used to sanction existing technical achievement, but will represent the goal towards which we must strive. (11)

Now, the distinction between primary and secondary ("feasible") standards is not unknown in the West. In the United States, for example, there is the traditional distinction between criteria--expressing the available scientific knowledge of the relationship between pollutants or toxic substances and their adverse effects on man and his environment--and prescriptive standards, i.e., norms established by some authority to govern action.(12)

Criteria are supposed to provide the scientific basis for the establishment of standards, and the two stages of the standard-setting process--the scientific and the regulatory--are sometime kept institutionally separate. Thus, the National Institute for Occupational Safety and Health (NIOSH), in the Department of Health, Education, and Welfare, has responsibility for developing criteria, while the Occupational Safety and Health Administration (OSHA), in the Department of Labor, sets standards guided by the criteria proposed by NIOSH.

However, the distinction between criteria and standards does not exactly correspond to that between primary and secondary standards. For example, NIOSH criteria are based not only on considerations of health and safety, but also on the feasibility of control within existing technology.(13) It has even been claimed that NIOSH criteria and recommendations

> are often politicized, a sphere reserved for OSHA; NIOSH is to arrive at the "best scientific judgment," objectively determined, on what constitutes safe exposures. However, in two cases NIOSH recommended criteria at variance with what was suggested by the scientific evidence alone. NIOSH recommended 90 dBA as an eight-hour noise exposure limit, eventually to go down to 85 dBA. Either level clearly causes a fair amount of hearing loss and takes no account of nonauditory effects. The NIOSH recommendation that the asbestos standard be set at 5 fibers per cc until 1976 ... is another example of how politicized the "objective" recommendations of NIOSH are... (14)

On the whole, environmental and occupational health legislation in the United States appears to be rather inconsistent on the role which non-health, particularly economic, considerations should play in the standard-setting process. Thus, while the Safe Drinking Water Act, the Toxic Substances Control Act, the Occupational Safety and Health Act, and the Federal Food, Drug, and Cosmetic Act (with the exception of the Food Additives Amendment of 1958) call for some weighing of the costs and benefits of regulation, the Clean Air Act Amendments, the Federal Water Pollution Control Act, and the Resource Conservation and Recovery Act are silent on this issue.

Even when the law requires some balancing of costs and benefits, the language is often ambiguous. In the case of the OSH Act, for example, the courts had to determine whether Section 6(b)(5) of the Act, which only speaks of feasibility ("The Secretary ... shall set the standard which most adequately assures, to the extent feasible, on the basis of the best available evidence, that no employee will suffer material impairment of health or functional capacity..."), permits OSHA to consider economic as well as technological factors in setting standards for toxic substances.

Since Congress has set few coherent guidelines on the extent to which benefits, as well as costs and risks, must be considered, regulators have to rely on ad hoc procedures to balance somehow these incommensurable factors. As a director of the Office of Toxic Substances of the Environmental Protection Agency has described the process,

> Usually the procedure is to postulate a numerical standard for a toxic chemical or a specific type of limitation on the use of the chemical, with the restriction designed to reduce environmental levels to the point that concerns over health or environmental damage disappear. Then an assessment is carried out to see if the favorable environmental impact from the restriction warrants the concomitant economic costs. If the costs are too high, the level of control is adjusted until an appropriate balance is reached.(15)

Because of the ambiguity of the legislative mandate, much current debate on health regulation has focused on whether the benefits and costs of specific health standards should be evaluated explicitly (possibly through a formal cost-benefit analysis) or implicitly, through bargaining and the political process. The outside observer cannot help being puzzled by this preoccupation with the fine-tuning of a mechanism--compulsory national standards--that seems to be intrinsically incapable of dealing with the myriads of local situations and problems that constitute the environmental or occupational-health problem in a country of the size of the United States. In the following pages I shall point out the limitations of compulsory standards as tools of public health policy, and discuss institutional conditions under which voluntary standards may provide a more adequte solution.

Before proceeding with the argument, however, it should be noted that even an unambiguous choice in favor of "health-only" criteria, as in the Soviet model, removes only some of the uncertainty and subjectivity that is inherent in the standard-setting process. The inadequacy of the scientific basis of regulation remains. For example, it has already been mentioned that the procedures used by toxicologists to determine "virtually safe doses" (VSD) for exposure to carcinogens involve extrapolations downward from the range of observed effects. Now, a variety of equally plausible mathematical functions may be used in the extrapolation procedure, but the choice of function has a major effect on the determination of the VSD--more than 100,000-fold according to estimates of the Advisory Committee on Safety Evaluation of the U.S. Food and Drug Administration.[16] Equally uncertain is the relevance of animal experiments for the determination of human carcinogenic risks.

In fact, scientific certainty in regulation is so elusive that, according to the suggestion of an experienced scientist,

> All things considered, it would seem reasonable that until better methods for the definition of relative toxicity can be found, the role of science in regulation should be limited to those instances where nearly certain assessment of human risk is feasible and legitimate; at the same time more emphasis should be given to methodological and basic research for future application.[17]

4. Standard-setting and standard-using

The formal process of adoption of health standards in the Soviet Union is quite straightforward. Government selects a scientific review committee that surveys the literature and exposure data, and makes recommendations to the Ministry of Health. If the Ministry agrees, these recommendations become nationwide regulations. Enforcement is the responsibility of the Ministry of Health (through the Sanitary Epidemiology Service, Sanepid), of inspectors from the All Union Councils (primarily engineers), and of inspectors from local trade union committees. Sanepid, with a staff of some 120,000 people, including 45,000 physicians, is responsible for both research and practice in preventive medicine and environmental and occupational health. Public health physicians, paramedical personnel (feldshers), and chemists in 4,500 Sanepid stations monitor pollutant levels, oversee the enforcement of standards, and participate in all health aspects of community planning.

But full enforcement of the existing standards is often impossible because the standards are numerous and in most cases very stringent. Hence, as already noted, temporary secondary standards, which take economic and technical constraints into consideration, are allowed. Conflicts between permitting a violation and closing down a plant appear to be fairly common, and are generally resolved through bargaining between Sanepid inspectors, local government, and the industry in question. According to observers from the West, the extreme step of closing a plant is resorted to infrequently, and only after a

period of arbitration involving the superiors of the local Sanepid inspector and the local industry manager.[18]

The apparent contradiction between the theory and the practice of standard-setting in the Soviet Union raises an interesting issue. One could argue that although their primary standards are typically much more stringent than those used in the West, actual differences tend to disappear at the level of enforcement. In more general terms, an approach combining stringent standards with "reasonable" enforcement could produce about the same results as a system of more realistic standards but stricter enforcement.

The proposition is plausible, but unfortunately no data are available to test it. And even if empirical testing was possible, one should be extremely careful in drawing conclusions. A result of "no significant difference," for example, would not tell us much about the long-run implications of the two approaches. Soviet regulators insist that even when their standards are not fully satisfied, "they represent the ultimate goal and enable us to assess, in each individual case, how far we have advanced in this difficult task."[19] They criticize standards used in the West for codifying existing economic and technical conditions, to the detriment of their normative character.

The point is well taken, if one assumes that the long-term effect of an environmental or occupational health standard should be to channel growth away from industries and materials that are hazardous to health and toward safer forms of production and employment. It is, in fact, hard to see how a "feasible"

standard can provide the necessary signals. Consider, for example, the history of OSHA regulation of the carcinogen vinyl chloride (VC).

In April, 1974, OSHA promulgated an Emergency Temporary Standard (ETS) reducing the previous National Consensus Standard from 500 parts per million (ppm) to 50 ppm. The National Consensus Standard of pre-OSHA times was a standard proposed by the American Conference of Governmental Industrial Hygienists, and voluntarily accepted by industry, at a time when it was unknown that VC could induce cancer. The statement of reasons supporting the ETS reveals that the 50 ppm standard was an uneasy compromise between conflicting considerations and interests.

During the summer of 1974, OSHA held extensive hearings on the proposal of a permanent standard of 1 ppm (as a time-weighted average over an 8-hour work period, with permissible excursions up to 5 ppm averaged over any 15 minute period). Although the disagreement on the medical evidence was considerable, most of the debate concerned the "feasibility" of the proposal. Industry opposed the proposed level of 1 ppm on the grounds that OSHA lacked sufficient evidence on the harmfulness of VC at low doses; that it was technologically impossible to meet the 1 ppm ceiling; and that the cost of approaching the ceiling would force the companies out of business. Conceding the industry's claims of infeasibility, OSHA finally promulgated a somewhat weakened permanent standard.

Subsequent experience was to show that meeting the 1 ppm level was not as difficult or costly as industry had predicted. The permanent VC standard lacks explicit criteria of feasibility, but a careful case study comes to the conclusion that "OSHA's statements and actions suggest that it was following an unarticulated principle that a standard is not feasible if it would cause more than slight changes in the number of firms in an industry, or in an industry's profit and growth rates, its output, and competitive position."(20)

The case of the vinyl chloride standard is far from being unique. American regulators are constantly urged to treat economic and technical feasibility as important considerations in the derivation of health standards. Responding to these pressures, regulators tend increasingly to conflate the conceptually distinct stages of standard-setting and standard-using. The resulting aggregation of scientific, technical, economic, and political criteria is not only ad hoc, but also logically inscrutable.(21) As a consequence, the meaning of the numerical value chosen for a given standard is ambiguous, representing neither a policy goal, nor a scientific judgment of health risk, nor even (since the standards are supposed to be enforceable at the national level) a measure of the level of protection that can be reasonably achieved in specific local situations. Whatever reservations one might have about the logic of the distinction drawn by Soviet regulators between primary and secondary standards, it must be admitted that at least it allows a clear statement of objectives, while avoiding the danger of sanctioning existing technical and economic conditions.

So far I have stressed the differences between the Soviet and the American approach. Despite these differences, both countries are now committed to a regulatory structure that relies on national mandatory standards; hence, both are faced by the same intrinsic limitations of this method of control. Mandatory standards focus the attention of operators and inspectors on a small set of permissible values and approved practices, at the expense of more comprehensive assessments of the overall quality of ambient or workplace environment. The logic of statutory control is such that it is difficult to differentiate between the important and the trivial, between form and substance. With no formal place for discretion in technical interpretation, the situation becomes one of either compliance or breach.

Moreover, given the limited knowledge available today in toxicology, radiation biology, epidemiology, and related fields, the numerical precision of current standards is spurious. At the same time, rigid statutory control does not allow the frequent revisions which a steady flow of new evidence would require. Nor can general regulations be written with enough specificity to accommodate all the unique conditions encountered in the millions of workplaces and thousands of communities of a large industrialized country.

Critics of the existing regulatory structure have used the inadequacies of mandatory standards as proof of the need to place greater reliance on economic incentives. It has been suggested, for example, that employers be induced to provide safer workplaces by means of an "injury tax," in the form of

a certain percent surcharge of an employer's total injury loss, which would raise the marginal benefits of injury prevention. This critical literature has its counterpart in the area of environmental problems. Here, too, the administrative approach to pollution control, based on standards and prohibitions, has been criticized for its lack of effectiveness, and for its tendency to become "a political process entailing bargaining between parties of unequal power."(22) Effluent charges and related price-based techniques have been proposed as alternative approaches that by their automatism "would reduce the scope for administrative discretion and bargaining."(23)

But these normative conclusions overlook one important point. The same forces that influence and distort the standard-setting process will also affect other approaches, perhaps by different means. The comparison between an uncorrupted system of effluent charges (or injury taxes), and a regulatory machinery captured by interest groups is a specious one. In fact, as I have shown elsewhere, where effluent charges have been used, e.g., in France, they have proved to be as subject to bargaining and as conditioned by the institutional framework as standards, licenses, and other administrative measures.(24)

There is no reason to believe that market-oriented approaches to occupational safety and health would fare better. At any rate, solutions that are (theoretically) more efficient will also be more desirable only to the extent that economic efficiency is accepted as the overriding criterion of public policy. In the area of environmental and occupational health such consensus on values seems to be lacking. Policy actors realize that the

choice between standards and prices is not a technical choice between policy tools that are in themselves neutral, but rather a decision between alternative institutional frameworks which reward different groups differently. Hence some actors may, without being inconsistent, recognize that standards are less efficient than economic incentives, and at the same time support an inefficient regulatory machinery in which they have a greater voice.

This, I would argue, is the basic reason for the support given by American labor unions, environmentalists, and other public interest groups to mandatory standards, set and implemented by the federal government. These groups have succeeded in bringing about significant changes from previous policies based on decentralized controls and voluntary compliance, but there is an ironic twist to their victory. Voluntary standards and guidelines, such as those used in pre-OSHA days, can be determined on the basis of "health-only" criteria since they are not meant to be regulatory instruments, but only to supply scientific inputs to subsequent decisions. Mandatory standards, on the other hand, <u>are</u> policy tools and as such must include, more or less explicitly, considerations of costs and benefits. To put it bluntly, one cannot object to the intrusion of economic and other non-health considerations in the regulatory process, to the trading of "lives for dollars," and at the same time insist on centralized statutory controls. Even the Soviets have found it necessary to operate with a dual set of standards, and it has already been suggested that their system may have some advantages over present American practices.

Fortunately, these are not the only possible alternatives. At least in the area of occupational safety and health, the West European model--in which standards are generally regarded as guidelines, inspectors have considerable discretionary power, and health and safety is regarded as the joint responsibility of management and labor--prefigures interesting possibilities of self-regulation that are absent in both the American and the Soviet system.

5. Self-regulation

One of the most significant characteristics differentiating the Western European from the American approach to occupational health and safety is the greater reliance of the former on voluntary, rather than legally enforceable, standards. Although extensive listings of maximum acceptable concentrations (MACs) for toxic substances and other environmental limits (many of them of American origin) are regularly published by nongovernmental and public research institutes, they typically represent nothing more than information to be used by the inspectors for the purpose of giving preventive advice and monitoring working conditions.

Good reasons for not embedding MAC values and other numerical standards in legal codes have been given by the Institut National de Securité of the French Ministry of Labor and Participation:

> We consider in effect that the individual reactions produced by a material are too unpredictable to permit fixing in a general fashion a margin of safety; that results obtained by animal experimentation are inapplicable as concerns the levels obtained by this manner; that the proposed criteria lacked comparability as to methods of

> investigation used, as well as biological response assayed and analytical methodologies, all of which are not uniform from one country to another or in the same country from laboratory to laboratory.(25)

In addition to a limited number of MAC values, the Institut National de Securité issues extensive "Fiches Toxicologiques" giving detailed information on physical, chemical, and toxicological characteristics of different substances. Private organizations, such as the Association Interprofessionelle des Centres Medicaux et Sociaux of the Paris region also publish their own Fiches Toxicologiques and extensive MAC listings, partly adapted from the lists of the American Conference of Governmental Industrial Hygienists (ACGIH).

In the Federal Republic of Germany basic research in occupational standards is done by the nongovernmental Commission for the Evaluation of Toxic Materials in the Workplace (Senatskommission zur Prüfung gesundheitsschädlicher Arbeitsstoffe). The Commission was created in 1955 by the Deutsche Forschungsgemeinschaft--the central organ of self-management of German scientific institutions--with the explicit goal of reducing excessive dependence on American standards. One of the important functions of the Commission is to provide scientific advice to regional and national parliaments and governments, and to local authorities. So far, the Commission has produced MAC listings concerning more than 400 substances. These MAC values are based only on health criteria; considerations of technical or economic feasibility are excluded.

German MAC values, like their French equivalents, are not embedded in legal codes, and can be modified and improved at

any moment. On the other hand, they may be made compulsory for the purpose of occupational health control simply by publishing them in Arbeitsschutz, a specialized journal published by the Federal Ministry of Labor and Social Security. Guidelines for the practical applications of the standards by federal and state factory inspectorates are issued by a Committee for Dangerous Materials in the Workplace (Ausschuß für gefährliche Arbeitsstoffe) set up by the Ministry of Labor, and whose membership includes representatives of the "Senatskommission." These guidelines interpret the MAC values in the light of existing technical and economic constraints.

In the United Kingdom, too, no specific legal status applies to occupational health standards, but values derived from the ACGIH threshold limit values, and from other sources, are adopted by the H. M. Factory Inspectorate of the Department of Employment for purposes of control and surveillance. In addition to serving as guides for administrative action, such values may be used in enforcement proceedings under the provisions of the Factory Act of 1961.

From the United Kingdom have come the most articulated proposals in recent years for a switch in emphasis away from an extensive use of statutory regulations toward greater reliance on voluntary standards and codes. I am referring to Safety and Health at Work, the official report of the Parliamentary Committee appointed in May, 1970 by the British Secretary of State for Employment and Productivity under the chairmanship of Lord Robens.[26] Among the recommendations made by the Robens

Committee the following are particularly relevant to the present discussion:

- Wherever practicable, regulations should be confined to statements of broad requirements in terms of the objectives to be achieved.

- In future, no statutory regulation should be made before detailed consideration has been given to whether objectives might adequately be met by a non-statutory code of practice or standard.

- Greater emphasis should be placed on standard-setting by means of non-statutory codes and standards. As a general rule, statutory regulations should only be made when the alternative of a non-statutory code or standard has been fully explored and found wanting.

- The whole regulatory system should be more flexibly based and more discriminating. The means used should encourage industry to deal with more of its own problems, thereby enabling official regulation to be more effectively concentrated on serious problems where strict official regulation is appropriate and necessary.[27]

These recommendations follow from the belief that statutory regulations are largely ineffective, intrinsically rigid, and have a built-in tendency to quickly become obsolete. On the other hand, "standards and codes developed within industry and by independent bodies are, over a large part of the field, more practical and therefore potentially more effective instruments of progress than statutory regulations."[28] The Report

concludes that what is needed is "less law" and more provision for voluntary self-regulation at the plant level.

However, in order to provide credible sanctions when needed, inspectors should have the power, without reference to the courts, to issue formal Improvement Notices, i.e., orders to comply not only with any relevant statutory regulation, but also with any relevant voluntary code or standard that has been formally approved by the Authority for Safety and Health at Work. Voluntary codes and standards would also be admissible evidence in proceedings before tribunals (the Report suggests that appeals against improvement notices should be heard not in the criminal courts but by the industrial tribunals set up under the 1964 Industrial Training Act). In cases where serious hazards or imminent dangers exist, the inspector could issue a Prohibition Notice ordering that, in the event of non-compliance within the stated time limit, the use of specified plant, machinery, processes or premises must be discontinued, or continued only under specific conditions.

But insuring compliance with minimum legal requirements is not the main task of the inspectorate. Rather, inspectors should be concerned with the broad aspects of safety and health organization at the workplace they visit, as much as with those narrow aspects which may have been made the subject of detailed statutory regulations. "We believe," the Report states, "that, as a matter of explicit policy, the provision of skilled and impartial advice and assistance should be the leading edge of the unified inspectorate."[29]

The provision of advice to industry and commerce, and the enforcement of sanctions where necessary, should be regarded as inseparable elements of inspection work under a policy that has as its prime objective the prevention of accidents and ill health and the promotion of progressively better standards at work. The success of such a policy requires close cooperation with the people who are exposed to the risks--the workers themselves.

> It should be as natural for inspectors to discuss safety and health problems with workpeople and their representatives as it is to discuss them with management... We are convinced that more contact and dialogue between inspectors and workpeople would not only greatly assist the inspectors in their day-to-day work, but would also make an invaluable contribution towards increasing the involvement of workpeople in the fight against safety and health hazards.(30)

The Robens Report raises a large number of regulatory issues that cannot be adequately dealt with in this paper. But at least one problem must be mentioned here, for it is of central importance from the viewpoint of comparative analysis. To what extent the recommendations of the Committee may be generalized to countries other than the United Kingdom? More generally, what are the functional requisites for a viable system of self-regulation and voluntary compliance? Three conditions seem to be essential: (1) A critical mass of highly qualified and uncorruptible inspectors; (2) The availability of significant penalties for serious violations; (3) An active concern on the part of management and workers and their representatives for the quality of the environment in the workplace.

In a number of countries of Western Europe these conditions are at least approximately satisfied.[31] In particular, issues of occupational health and safety play an increasingly important role in collective bargaining and trade union policies. In France and Italy, for example, unions of different ideological persuasions have consistently refused, in recent years, to trade off health risks with higher wages. They have also been very active in informing their members about such risks and in mobilizing public support for preventive measures. Their example is now being followed in other European countries, as the following episode illustrates. In 1980 the Swiss Union of Textile, Chemical, and Paper Workers (GTCP) issued a well-written popular report (Gesundheit am Arbeitsplatz = Health at the Workplace) providing information about the major problems of occupational health and stress, and outlining a model for workers' monitoring and control of health problems at the workplace. This includes the systematic collection of environmental and health data at the plant level by means of report cards to be mailed directly to union headquarters in Zürich for central evaluation. Interestingly enough Gesundheit am Arbeitsplatz is the German translation of an analogous document prepared by the Italian Union of Metal Workers--a politically much more radical organization than its Swiss counterpart.

Some American analysts doubt that a system of self-regulation would work in the United States. American inspectors do not share the prestige and long tradition of their European colleagues, and also their training is apparently not as good. The pre-OSHA experience with "consensus standards" voluntarily

adopted by industry under lax supervision by the states has been sharply criticized by labor unions and public interest groups. In fact, national labor organizations have been among the most determined supporters of compulsory federal regulation.(32)

It is probably true that passage of the OSH Act has initially strengthened the position of the union leadership in relation to management, to the rank and file and their local representatives (to whom financial gains are of more immediate interest than improved working conditions), and to the unorganized member of the workforce. Yet, even the most ardent supporters of federal regulation cannot by now fail to see the intrinsic limitations of the present regulatory structure. That OSHA has managed to produce only ten health standards in ten years is something that cannot be explained in terms of incompetence or poor management, but has its roots in basic ambiguities in regulatory philosophy, and in a poor choice of policy tools.

The acute dissatisfaction with OSHA's standard-setting and enforcement activities has elicited a number of suggestions for how policy might be improved. Among the alternative strategies that have been proposed, regulation through collective bargaining is perhaps the most promising. In the words of a recent writer,

> ...it may make sense to involve unions in abating hazards in the workplace, rather than relying exclusively on governmental regulation and inspection... To the extent that labor and management can be induced to negotiate health and safety rules within the context of the collective bargaining agreement, we can decentralize some aspects of the regulatory intervention mechanisms almost to the plant level. In theory, this should produce investments in occupational health and safety, that are both more efficient and more effective than those produced by the present system alone.(33)

I would like to conclude by suggesting that the rich experience of the European labor movement in the area of occupational health and safety could be of considerable help in designing a strategy of regulation through collective bargaining adapted to the American context. For this reason that experience deserves to be carefully studied by American analysts.

Notes

1. World Health Organization, Technical Report Series 647, Geneva: 1980, p. 13.
2. Theodore F. Hatch, "Permissible levels of exposure to hazardous agents in industry," Journal of Occupational Medicine, vol. 14, 1972, pp. 134-137.
3. Ib., p. 135
4. Roger I. Glass, "A perspective on environmental health in in the USSR," Archives of Environmental Health, vol. 30, August 1975, pp. 391-395.
5. V. A. Rjazanov, "Criteria and methods for establishing maximum permissible concentrations of air pollution," Bulletin of the World Health Organization, 32 (1965), pp. 389-398, at p. 392.
6. Bertram D. Dinman, "Development of workplace environment standards in foreign countries, Part 1," Journal of Occupational Medicine, vol. 18, no. 6, 1976, pp. 409-417.
7. Mark G. Field, Soviet Socialized Medicine, New York: The Free Press, 1967, ch. 9.
8. B. D. Dinman, cit.
9. R. I. Glass, cit.
10. V. A. Rjazanov, cit., p. 390.
11. Ib., p. 390

12. See, for example, William W. Lowrance, Of Acceptable Risk, Los Altos, Calif.: William Kaufmann, Inc., 1976, pp. 133-136.

13. Charles H. Powell and Herbert E. Christensen, "Development of Occupational Standards," Archives of Environmental Health, vol. 30, April 1975, pp. 171-173.

14. Nicholas A. Ashford, Crisis in the Workplace, Cambridge, Mass.: The MIT Press, 1976, p. 301.

15. Glenn E. Schweitzer, "Toxic Chemicals and Regulatory Decision Making: Philosophy and Practicality," Appendix B of National Academy of Sciences, Decision Making for Regulating Chemicals in the Environment, Washington, D. C.: 1975, pp. 71-78, at pp. 72-73.

16. Jerome Cornfield, "Carcinogenic Risk Assessment," Science, vol. 198, 18 November 1977, pp. 693-699.

17. Gio Batta Gori, "The regulation of carcinogenic hazards," Science, vol. 208, 18 April 1980, pp. 256-261, at p. 260.

18. Roger I. Glass, "A perspective on environmental health in the USSR," cit., p. 393.

19. V. A. Rjazanov, "Criteria and methods for establishing maximum permissible concentrations of air pollution," cit., p. 390.

20. David D. Doniger, The Law and Policy of Toxic Substances Control, Baltimore: The Johns Hopkins University Press, 1978, p. 65.

21. See, for example, the comments by G. E. Schweitzer, in "Toxic Chemicals and Regulatory Decision Making," cit., especially pp. 72-73.

22. A. M. Freeman, R. H. Haveman, A. V. Kneese, The Economics of Environmental Policy, New York: Wiley and Sons, 1973, p. 105.

23. Ib., p. 170.

24. Giandomenico Majone, "Choice among policy instruments for pollution control," Policy Analysis, vol. 2, no. 4, 1976, pp. 589-613.

25. Cited by Dinman, "Development of workplace environment standards in foreign countries, Part 1," cit., p. 411.

26. Safety and Health at Work, Report of the Committee 1970-72, Chairman Lord Robens, London: H. M. Stationary Office, Cmnd. 5034, 1972.

27. Ib., pp. 44-46.

28. Ib., p. 48.

29. Ib., p. 65.

30. Ib., p. 65.

31. See, for example, N. A. Ashford, Crisis in the Workplace, cit., pp. 501-510.

32. John Mendeloff, Regulating Safety, Cambridge, Mass.: The MIT Press, 1979, p. 16.

33. Lawrence S. Bacow, Bargaining for Job Safety and Health, Cambridge, Mass.: The MIT Press, 1980, p. 58.

APPENDIX

The Scientific Basis of Regulation

This appendix elaborates some technical points to which reference has been made in sections 2-4 of the paper. The first issue refers to the nature of the dose-response function and the procedures used to extrapolate it outside the range of observed responses.

Two main methods have been developed in order to determine an acceptable level of risk corresponding to a given exposure to a toxic substance. The first one involves the notion of a "no observed effect level" (NOEL), and is the standard toxicological procedure used in the United States and many other countries. The NOEL is defined as the quantity of a substance administered to a group of experimental animals at which effects observed at high levels are absent, and at which no significant differences between the group of animals exposed to that quantity and the control group are produced. The acceptable level of risk is obtained by dividing the NOEL by 100. This particular value of the safety factor is commonly justified by the rule of thumb that man may be ten times more sensitive that the experimental animals used, and that there may be in addition a tenfold variation in sensitivity among individual animals.

For testing carcinogenic substances a second, more refined approach has been developed. Its characteristic feature is the use of an explicit mathematical model expressing the probability of a lifetime response as a function of dosage $D : P = f(D)$. Depending on the choice of the response function f, different models arise.

In the following discussion a quantal response will be assumed. This means that each experimental animal is assumed to be characterized by a tolerance, so that any dose above it will induce cancer and any below it will not. It is then natural to consider the distribution of tolerances to different dose levels over the experimental population. Thus f(D) becomes the density of the probability distribution of responses, and the proportion of the population that will respond to a dose level D_0 is given be the cumulative function

$$P(D_0) = \int_0^{D_0} f(D)dD \quad .$$

In particular $P(0) = 0$ (i.e., there is no spontaneous occurrence of the particular response), and $P(\infty) = 1$, (i.e., no immune group exist within the population; all members will respond to sufficiently high doses). It should be noted that not all toxicologists agree with these particular implications of the model.

A virtually safe dose (VSD) D_0 is a dose level such that $P(D_0) \leq P_0$, where P_0 is some preassigned small probability such as 10^{-8} or 10^{-6}, the latter value being the one used, for example, by the U.S. Food and Drug Administration. The VSD can be computed as $P^{-1}(P_0)$, as soon as f is known.

How is the function f determined? What the experimenter actually observes is a sequence of dose levels D_i, $i = 1,2,\ldots,n$ and the corresponding proportion of responses $p_i = P(D_i)$. The n observed pairs (D_i,p_i) can be used to interpolate the dose-response function; outside the observed range responses are estimated by extrapolation.

Toxicity tests often show that the observed proportion of responses increases monotonically with dose and exhibits a sigmoid shape. This still allows many degrees of freedom in choosing the functional form for f(D). One often-used model postulates that the distribution of responses in the population is given by a normal probability density:

$$P(D) = \int_{-\infty}^{\frac{\chi(D)-\theta_1}{\theta_2}} (2\pi)^{-1/2} \exp\left(-\frac{1}{2} u^2\right) du$$

where θ_1 and θ_2 are parameters to be estimated from the experimental observations according to the maximum likelihood or some other statistical method, and $\chi(D)$ is some transformation of the dose level D, e.g., $\chi = \log(D)$.

The argument leading to the normal (probit) function is a purely statistical one. Because of inherent biological variability not all animals will have the same tolerance. Hence, assuming a quantal response, at any given dose level D only animals with tolerance below D will respond. If the distribution of responses is assumed to be log-normal, then the probit-log dose function results.

A different line of reasoning, this time suggested by chemical kinetic theory, leads to another functional form, namely the logistic curve

$$P(D) = \frac{D^{\theta_2}}{\theta_1 + D^{\theta_2}}$$

where θ_1 and θ_2 are again parameters to be estimated.

Still other dose-response models have been proposed on the basis of the so-called "hit theory" of carcinogenicity. The "one hit" hypothesis assumes that cancer is an expression of a permanent change in cellular genetics resulting from the interaction of one molecule of carcinogen with a critical receptor in one cell. Under the "multi-hit" hypothesis two or more events are necessary for induction of a tumor. The single-hit model leads to

$$P(D) = 1 - \exp(-\theta D) \quad , \theta > 0$$

while the multi-hit model suggests the function

$$P(D) = 1 - \sum_{k=0}^{m} (\theta D)^k \exp(-\theta D) \quad , \theta > 0$$

where m is the minimum number of hits on a receptor required to induce a response. More complicated expressions for the dose-response relationship have also been suggested.

From the viewpoint of the regulator this profusion of models is a source of embarassement rather than enlightenment since (a) the different dose-response functions often cannot be distinguished from each other in the observable range; (b) no firm scientific basis at present exists for choosing among them; and (c) the choice of function has a major effect on the VSD and hence on the level of risk that is considered to be acceptable.

The following table, taken from the report of the Advisory Committee on Protocols for Safety Evaluation of the U.S. Food and Drug Administration,[1] shows the widely different results given by three of the most frequently used models at low doses. It should

be noted that these three models are hardly distinguishable in the observed range of 5% - 95% response rates.

Dose level	Log-normal model	Log-logistic model	Single-hit model
0.01	0.05	0.4	0.7
0.001	0.00035	0.026	0.07
0.0001	0.0000001	0.0016	0.007

It may be argued that if there is no firm scientific basis for choosing among the different mathematical models, then one should prefer the safest or most conservative procedure. One problem with the conservatism argument, as Jerome Cornfield points out,[2] is that it is not clear where one should stop. A no-threshold model is more conservative than one that admits the existence of thresholds for carcinogenic effects. But within the large class of no-threshold models many degrees of conservatism are possible. Again, in designing a toxicological experiment one could use the most sensitive species, the most sensitive strain within species, and so on down to the level of the most sensitive individual animal, thus obtaining 100 percent incidence at each dose level. In short, it is difficult to be conservative in a consistent manner, unless one is prepared to propose a zero level of exposure in each case.

As a decision rule, conservatism in the face of risk is as unsatisfactory as the minimax or the "most likely event" principles, or indeed any principle that does not balance expected risks against expected benefits. On the other hand, the only consistent (Bayesian) decision procedure requires information inputs--prior probabilities of all possible scien-

tific hypotheses, utilities for all possible consequences--which no regulator is likely to supply.

What about determining acceptable levels of exposure on the basis of a cost-benefit or a risk-benefit analysis? There are, of course, well-known difficulties in quantifying benefits, costs, and risks. The danger that the estimates represent little more than disguised value judgments is always present. Problems of quantification aside, a number of rather stringent conditions must be satisfied before the cost-benefit criterion may be meaningfully used as a decision rule in health regulation.

First, it has long been recognized that cost-benefit analysis is not applicable under extreme circumstances, for example, when potential health damage is so large that marginal trade-offs between the risks and the benefits of certain kinds of activity become virtually meaningless. Also, price-based marginal calculations (and the *tâtonnement* procedures necessary to discover a correct set of prices) are hardly appropriate when immediate action is required. To use Martin Weitzman's example, suppose that a certain number of airplanes is required for an emergency operation. In strict economic terms, it would be inefficient to issue orders to different commercial airlines to supply a given number of airplanes, since marginal opportunity costs will typically vary from company to company. Yet, in practical terms, this approach would be preferable to the economically correct procedure of announcing a price for plane services and letting profit-maximizing companies decide on the number of planes they would be willing to commit to the rescue operation.[3]

A third case in which the relevance of the approach is doubtful is when the biological effect of the toxicant to be regulated becomes evident only after a long time (perhaps 20 to 30 years). If people are unaware of such long-run effects, no externalities are generated and the cost-benefit criterion would indicate a status quo policy--a solution most toxicologists would consider irresponsible. Moreover, persistent pollutants (like cadmium, mercury, or radioactive wastes, which deteriorate very slowly over time) pose particular problems, since the damage they cause arises mostly from a non-reducible *stock*, while only incremental damage is caused by the *flow* of pollution. Now, it is a basic assumption of the marginal calculus that the relevant variables can be controlled in all directions. In the case of persistent pollutants the stock of pollution is, to all practical purposes, non-reducible, so that an essential feature of the cost-benefit approach is missing in this toxicologically important situation.

However, the most serious difficulties in using cost-benefit criteria arise in connection with dynamic aspects of pollutant exposure. The following considerations, which closely follow an argument presented by David Pearce in a slightly different context,[4] show that even if the threshold hypothesis is accepted, cost-benefit ratios are not reliable guides toward a dynamically stable situation. In Figure 1 E(x) denotes the amount of pollutant exposure, as a function of the level of output or consumption of a substance x, while T_0 is the threshold level for a representative individual with respect to the given pollutant. We assume $E'(x) > 0$. B(x) is the marginal net private benefit, while $C_0(x)$ represents marginal external costs caused by the pollutant above threshold level T_0.

Now, given the assumption of a safe threshold at T_0, the external-costs function $C_0(x)$ has its origin at x_0, where $E(x)$ intersects T_0. Cost-benefit analysis indicates x_p as the Pareto-optimum level of output or consumption, since at x_p marginal net private benefits equal external costs. But at x_p, $E(x_p) > T_0$, and this implies that the situation may be biologically unstable, for the ability of the organism to cope successfully with the insult is now impaired. Unless we assume that full biochemical adaptation takes place, the limit of tolerance for subsequent exposure must be set at some lower level T_1. The same argument can now be repeated with respect to T_1 and a new external-costs function $C_1(x)$. The new Pareto-optimal level is $x_{p'}$, and since $E(x_{p'}) > T_1$, a standard set according to the cost-benefit criterion would lead to a further deterioration of the biochemical adaptability of the individual. In turn, this would cause a further reduction of the threshold level, and so on. Zero output or consumption (or, alternatively, total destruction of the ability of the organism to adapt following pollutant exposure) is the limiting solution of the sequence of such Paretian adjustments. To avoid this, it is necessary to fix the standard in correspondence to x_0 from the very beginning and this is what regulators attempt to do in practice. What our argument shows is that a stable standard could not be derived from cost-benefit consideration, since these fail to take into account the impact of an allegedly optimal solution on the dynamics of biochemical adaption. Hence, the cost-benefit approach can have only limited applicability in the management of serious biological hazards.

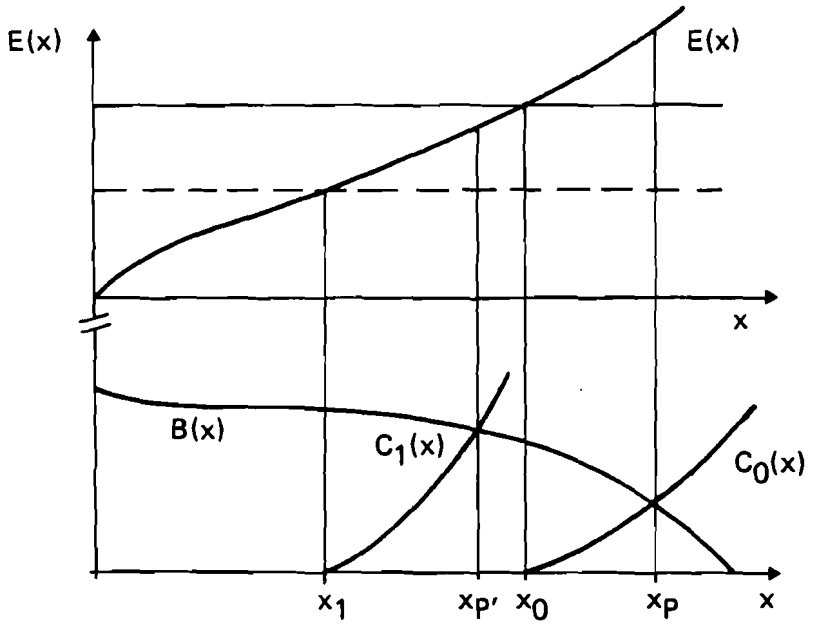

Figure 1: Threshold Values and Cost-Benefit Analysis (Source: Adapted from D. Pearce, *Kyklos*, 1976).

In sum, in the present state of knowledge risk assessment must be regarded as a trans-scientific activity (in Alvin Weinberg's sense) since it involves questions that can be stated in scientific language but are beyond the capacity of science to answer. At least for the time being, regulatory decisions must be based on prudential reasoning rather than on mathematical or logical formalisms.

NOTES

1. Food and Drug Administration Advisory Committee on Protocols for Safety Evaluation, "Panel on carcinogenesis report on cancer testing in the safety evaluation of food additives and pesticides," Toxicol. Appl. Pharmacol., 20, 1971, pp. 419-438.

2. Jerome Cornfield, "Carcinogenic Risk Assessment," Science, vol. 198, 18 November 1977, pp. 693-699.

3. Martin L. Weitzman, "Prices vs. Quantities," Review of Economic Studies, 41, October 1974, pp. 477-491.

4. David Pearce, "The Limits of Cost-Benefit Analysis as a Guide to Environmental Policy," Kyklos, vol. 29, Fasc. 1, 1976, pp. 97-112.