Chapter 8

ENVIRONMENTAL LAW

RICHARD L. REVESZ*
School of Law, New York University

ROBERT N. STAVINS†
John F. Kennedy School of Government, Harvard University, and Resources for the Future

Contents

1. Introduction 502
2. Setting goals and targets: the ends of environmental policy 504
   2.1. Normative issues and analysis 505
      2.1.1. Criteria for environmental policy evaluation 505
      2.1.2. Benefit–cost analysis of environmental regulations 508
      2.1.3. Other approaches to analyzing the goals of environmental policies 524
   2.2. Positive issues and analysis 526
      2.2.1. Executive orders 526
      2.2.2. Legislative enactments 528
      2.2.3. Judicial recognition 533
      2.2.4. A political economy perspective on how standards are set 533
3. Choosing instruments: the means of environmental policy 534
   3.1. Normative issues and analysis 535
      3.1.1. Potential criteria for choosing among policy instruments 535
      3.1.2. Alternative policy instruments 536
      3.1.3. Cross-cutting issues 552
      3.1.4. Normative lessons 556
   3.2. Positive issues and analysis 558
      3.2.1. Historical dominance of command-and-control 558
      3.2.2. Prevalence of tradeable permits allocated without charge 561
      3.2.3. Increased attention to market-based instruments 562

* Revesz is Dean of the New York University School of Law, where he is the Lawrence King Professor of Law, and Director, Program on Environmental Regulation, Center on Environmental and Land Use Law.
† Stavins is the Albert Pratt Professor of Business and Government at the John F. Kennedy School of Government, Director of the Harvard Environmental Economics Program, and a University Fellow, Resources for the Future.

Handbook of Law and Economics, Volume 1
Edited by A. Mitchell Polinsky and Steven Shavell
© 2007 Elsevier B.V. All rights reserved
DOI: 10.1016/S1574-0730(07)01008-0
4. Allocation of responsibility across levels of government

4.1. Positive review of responsibility of levels of government

4.2. Normative review of allocation of regulatory responsibility

4.2.1. Competition among political jurisdictions: the race to the bottom

4.2.2. Transboundary environmental problems

4.2.3. Public choice and systematic bias

5. Conclusions

Acknowledgements

References

Abstract

This chapter provides an economic perspective of environmental law and policy. We examine the ends of environmental policy, that is, the setting of goals and targets, beginning with normative issues, notably the Kaldor–Hicks criterion and the related method of assessment known as benefit–cost analysis. We examine this analytical method in detail, including its theoretical foundations and empirical methods of estimation of compliance costs and environmental benefits. We review critiques of benefit–cost analysis, and examine alternative approaches to analyzing the goals of environmental policies.

We examine the means of environmental policy, that is, the choice of specific policy instruments, beginning with an examination of potential criteria for assessing alternative instruments, with particular focus on cost-effectiveness. The theoretical foundations and experiential highlights of individual instruments are reviewed, including conventional, command-and-control mechanisms, market-based instruments, and liability rules. Three cross-cutting issues receive attention: uncertainty; technological change; and distributional considerations. We identify normative lessons in regard to design, implementation, and the identification of new applications, and we examine positive issues: the historical dominance of command-and-control; the prevalence in new proposals of tradeable permits allocated without charge; and the relatively recent increase in attention given to market-based instruments.

We also examine the question of how environmental responsibility is and should be allocated among the various levels of government. We provide a positive review of the responsibilities of Federal, state, and local levels of government in the environmental realm, plus a normative assessment of this allocation of regulatory responsibility. We focus on three arguments that have been made for Federal environmental regulation: competition among political jurisdictions and the race to the bottom; transboundary environmental problems; and public choice and systematic bias.
Keywords

environmental economics, environmental law, efficiency, cost-effectiveness, benefit–cost analysis, environmental federalism

JEL classification: K320, Q280, Q380, Q480
1. Introduction

An economic perspective can provide clarity regarding the causes and consequences of environmental degradation, and thereby provide insights regarding public policies intended to protect the environment. This is true both with regard to normative and positive assessments of environmental policies. Despite this value, an economic perspective is by no means a perfect substitute for other legitimate perspectives on environmental law and policy, whether from the natural sciences, from ethics, or from other disciplines. Rather, an economic perspective is a valuable complement to such views. Indeed, over the past several decades, as the attention given to environmental issues in the United States has grown, greater consideration has also been given to the efficiency, cost-effectiveness, and distributional equity of laws and regulations intended to protect the environment.1

In an effort to be rigorous in our review while keeping the treatment to reasonable length, we have imposed limits on the scope of our coverage. First, we focus on pollution control, and do not consider natural resource management, despite the fact that these two areas are closely related. Second, we concentrate our attention on environmental protection efforts at the federal level in the United States, and do not examine state, local, or international regulatory efforts.

We begin with the core question of whether and why environmental regulation is needed, considering the fact that under many conditions unconstrained markets produce socially desirable outcomes. What about in the environmental sphere? Under what specific circumstances will governmental intervention be appropriate? The fundamental theoretical argument for government activity in the environmental realm is that pollution is a classic example of an externality—an unintended consequence of market decisions, which affects individuals other than the decision maker. Because firm-level decisions do not take into account full social costs, pollutant emissions tend to be higher than socially efficient levels. As environmental quality is thus naturally under-provided by competitive markets, a possible role arises for government regulation. The traditional theoretical solution to the externality problem was long thought to be to force private actors to “internalize” the full costs of their actions. The primary advocate of this view was Arthur Pigou, who in *The Economics of Welfare* (1920) proposed that the government should impose a tax on emissions equal to the cost of the related damages at the efficient level of control.

A critical response to the Pigovian perspective was provided by Ronald Coase in his seminal article, *The Problem of Social Cost* (1960). Coase made three key points.1

---

1 We follow the standard definition of an **efficient** environmental policy as being one that involves a target—such as a 50 percent reduction in sulfur dioxide (SO2) emissions—that maximizes the difference between social benefits and social costs, i.e., a target level at which marginal benefits and marginal costs are equated. By **cost-effective** policies, we refer to those that take (possibly inefficient) targets as given by the political process, but achieve those targets with policy instruments—such as a tradeable permit system in the SO2 case—that minimize aggregate costs. Assessments of the **distributional** implications of environmental policies include analyses of the incidence of costs and of benefits.
First, he argued that even if A’s actions inflict harm on B, it is not the case that A’s actions should necessarily be restrained, because A’s harm of B is really “a problem of a reciprocal nature” which arises because of the simultaneous presence of two parties. For example, the problem of a factory that emits fumes that harm a nearby laundry is not caused solely by the factory. Protecting the laundry by enjoining the fumes would impose harm on the factory, just as protecting the factory by not enjoining its actions would impose harm on the laundry.

Second, Coase demonstrated that in a bargaining environment without transaction costs, parties will reach socially desirable agreements; and third, that the overall amount of pollution will be independent of the legal rules (assignment of property rights) chosen to structure their relationship. For example, if the legal regime enjoined pollution, but the harm to the factory were greater than the harm that the laundry would have suffered in the absence of such an injunction, the parties will enter into a contract under which, in return for a payment, the laundry will agree not to exercise its right to seek an injunction. Conversely, if the legal regime allows the pollution but the resulting harm to the laundry is greater than the harm that the injunction would impose on the factory, the parties will enter into a contract under which, again in return for a payment, the factory would agree not to pollute. Thus, regardless of the initial legal rule, bargaining will produce two results: (1) it will lead to the same amount of pollution; and (2) it will lead to the maximization of social welfare. Of course, the choice of legal rules can determine which party makes payments and which party receives them, a distributional concern, though not one of efficiency.

These three points are jointly characterized as the Coase Theorem. The Theorem may be said to hold if there are no transaction costs, no wealth or income effects, private rather than public goods, and no third-party impacts (i.e., all affected parties participate in the negotiation). At least some of these conditions are unlikely to hold in the case of most environmental problems. Hence, private negotiation will not—in general—fully internalize environmental externalities. And when market transactions—including Coasian bargaining—do not generate socially efficient allocations of resources, government regulation may be necessary to improve environmental quality.

On the other hand, although government regulation may be necessary to improve environmental quality when market transactions fail to generate socially efficient allocations of resources, such regulation is by no means sufficient to improve welfare or even environmental quality. This is because government regulation itself may not be efficient, that is, government may under-regulate or over-regulate, and/or it may regulate in ways that require unnecessarily large costs of compliance.

---

2 Coase recognized that transaction costs can be significant, and could prevent efficient negotiated outcomes. When transaction costs are great, the choice of legal rule will affect the amount of pollution and hence the level of social welfare.

3 That is, when the size of the payments is sufficiently small relative to the firms’ or individuals’ incomes or wealth that payment and receipt has no effect on respective supply and demand functions.

4 The public choice literature in economics suggests specific reasons for “government failure,” analogous to market failure. See our application of the economic theory of politics in section 3.2, for example.
We continue in section 2 of this chapter with an examination of the ends of environmental policy, that is, the setting of goals and targets. We begin with an examination of normative issues, notably the Kaldor–Hicks criterion and the related method of benefit–cost analysis. We examine this analytical method in detail, including its theoretical foundations and various approaches to the estimation of compliance costs and environmental benefits. We include a review of critiques of benefit–cost analysis, and examine alternative approaches to analyzing the goals of environmental policies. The section closes with a positive survey of the efforts of the Federal governmental to use these analytical methods.

In section 3, we turn to the means of environmental policy, that is, the choice of specific policy instruments. We begin with normative issues, and examine potential criteria for assessing alternative instruments, with particular focus on cost-effectiveness. The theoretical foundations and experiential highlights of individual instruments are reviewed, beginning with conventional, command-and-control, and then turning to economic incentive or market-based instruments. In the latter category, we consider pollution charges, tradeable permit systems, market friction reductions, and government subsidy reductions. We also consider the role of liability rules. Three cross-cutting issues merit particular attention: implications of uncertainty for instrument choice; effects of instrument choice on technological change; and distributional considerations. From this review, we identify a set of normative lessons in regard to design, implementation, and the identification of new applications. The section closes with an examination of positive issues, including three phenomena we seek to explain: the historical dominance of command-and-control; the prevalence in new proposals of tradeable permits allocated without charge; and the relatively recent increase in attention given to market-based instruments.

In section 4, we turn to the question of how environmental responsibility is and should be allocated among the levels of government. We offer a positive review of the responsibilities of Federal, state, and local levels of government in the environmental realm, plus a normative assessment of this allocation of regulatory responsibility. We examine three arguments that have been made for federal environmental regulation: competition among political jurisdictions and the “race to the bottom;” transboundary environmental problems; and public choice problems. In section 5, we conclude.

2. Setting goals and targets: the ends of environmental policy

If it is deemed appropriate for government to become involved in environmental protection, how intensive should that activity be? In real-world environmental policy, this question becomes how stringent should our environmental goals and targets be? For example, should we cut back sulfur dioxide (SO₂) emissions by 10 million tons, or would a 12 million ton reduction be better? In general, how clean is clean enough? How safe is safe enough?
2.1. Normative issues and analysis

Most economists would argue that economic efficiency—achieved when the difference between benefits and costs is maximized—ought to be one of the fundamental criteria for evaluating environmental protection efforts. Because society has limited resources to spend, benefit–cost analysis can help illuminate the trade-offs involved in making different kinds of social investments. In practice, there are significant challenges, in large part because of inherent difficulties in measuring benefits and costs. In addition, concerns about fairness and process merit consideration, because public policies inevitably involve winners and losers, even when aggregate benefits exceed aggregate costs.

2.1.1. Criteria for environmental policy evaluation

More than 100 years ago, Vilfredo Pareto enunciated the well-known normative criterion for judging whether a social change—possibly induced by public policy—makes the world better off: a change is **Pareto efficient** if at least one person is made better off, and no one is made worse off (1896). This criterion has considerable normative appeal, but virtually no public policies meet the test of being true Pareto improvements, since there are inevitably some in society who are made worse off by any conceivable change. Nearly fifty years later, Nicholas Kaldor (1939) and John Hicks (1939) postulated a more pragmatic criterion that seeks to identify “potential Pareto improvements:” a change is defined as welfare-improving if those who gain from the change could—in principle—fully compensate the losers, with (at least) one gainer still being better off.

The Kaldor–Hicks criterion—a test of whether total social benefits exceed total social costs—is the theoretical foundation for the use of the analytical device known as benefit–cost (or net present value) analysis. Neither the Pareto efficiency criterion nor the Kaldor–Hicks criterion calls for support for any policy for which benefits are greater than costs. Rather, the key is to identify the policy for which the positive difference between benefits and costs is greatest; otherwise it would be possible to identify another policy that would represent a further (potential) Pareto improvement.

If the objective is to maximize the difference between benefits and costs (net benefits), then the related level of environmental protection (pollution abatement) is defined as the efficient level of protection:

\[
\max_{\{q_i\}} \sum_{i=1}^{N} \left[ B_i(q_i) - C_i(q_i) \right] \rightarrow q_i^* \tag{1}
\]

where \( q_i \) is abatement by source \( i \) (\( i = 1 \) to \( N \)), \( B_i(\cdot) \) is the benefit function for source \( i \), \( C_i(\cdot) \) is the cost function for the source, and \( q_i^* \) is the efficient level of protection (pollution abatement). The key necessary condition that emerges from the maximization problem of equation (1) is that marginal benefits be equated with marginal costs.
The Kaldor–Hicks criterion is clearly more practical than the strict Pareto criterion, but its normative standing is less solid and has been attacked from various quarters. Although basic economic (utility) theory posits that individual well-being is a function of the satisfaction of individual preferences, this notion has been debated in other disciplines, including psychology and philosophy. In addition, questions have been raised about whether social gains and losses can be expressed through the simple aggregation of welfare changes of individuals. Some have argued that other factors should be considered in a measure of social well-being, and that criteria such as distributional equity should trump efficiency considerations in some collective decisions (Kelman, 1981a; Sagoff, 1993). Many economists do not disagree with this assertion, and indeed have noted that the Kaldor–Hicks criterion should be considered neither a necessary nor a sufficient condition for public policy (Arrow et al., 1996b).

At the heart of the claim that the Kaldor–Hicks criterion lacks normative standing for public decision making is the lack of any guarantee that compensation can or will be paid. Gains and losses to individuals can be aggregated in a variety of ways, but the standard method of aggregation is a simple sum, an approach that can be problematic if there is diminishing marginal utility of income or individual utility is dependent on the overall societal distribution of income. Under such conditions, policies can pass a potential Pareto improvement test, but decrease overall societal well-being, or vice-versa. Thus, some of the debate may be understood as focusing on the compatibility of the efficiency and distributional equity criteria. The general view from economics is that other criteria in addition to efficiency can and should be employed by policy makers, but that the existence of such criteria does not invalidate the efficiency criterion, which should remain part of social decision-making (Arrow et al., 1996b; Kopp, Krupnick, and Toman, 1997).

Many proposed and implemented environmental policies involve real trade-offs between equity and efficiency, and both international and national policy bodies have demonstrated concern for ensuring that groups such as low-income citizens, ethnic minorities, and future generations do not bear “disproportionate” shares of the costs of

---

5 If a proposal fails the (weaker) Kaldor–Hicks test, it cannot pass the Pareto test. Hence, at a minimum, the Kaldor–Hicks criterion can be used to weed out the worst policies, that is, those that cannot make the world better off in the Pareto efficiency sense.


7 For a contrasting view, see Kaplow and Shavell (2001, 2002a), who argue that any policy assessment that accords importance to non-utility criteria violates the Pareto principle and, thus, is subject to powerful criticism. See the discussion in section 3.1.3.3, below, of distributional considerations.


9 Data limitations sometimes reduce the reliability of economic benefit estimates, thus reducing the efficacy of benefit–cost analysis and the operational content of the efficiency criterion. Economics can still aid in decision making through the cost-effectiveness criterion, where an environmental target is taken as given, and the least-cost means of achieving that target are identified. We consider cost-effectiveness analysis later, in the context of normative analysis of policy instrument choice.
environmentally related actions.\(^{10}\) While it is conceivable to combine the goals of equity and efficiency using a social welfare function to arrive at a single metric (Bergson, 1938; Jorgenson, 1997), the information constraints and collective choice caveats have been acknowledged (Arrow, 1963, 1977; Sen, 1970). The consensus, at least within the realm of environmental policy, is that efficiency and equity ought to be evaluated separately (U.S. Environmental Protection Agency, 2000a), but there is no consensus on specific criteria that might be used to rank alternatives from an equity perspective.\(^{11}\)

In recent years, there has been much debate among economists, and between economists and nearly everyone else regarding the meaning of the frequently employed concept of “sustainability.” Ecologists and many others outside the economics profession have taken sustainability to be the unique and comprehensive criterion that can and should guide global development. In contrast, economists have tended to define sustainability as being purely about intertemporal distribution, that is, intergenerational equity.\(^{12}\) As such, most economists have viewed sustainability as no more than one element of a desirable development path.

A broader notion of sustainability, with considerable appeal outside of economics, combines two components—dynamic efficiency and intergenerational equity (Stavins, Wagner, and Wagner, 2003). Thus, a sustainable path is one that is both efficient and non-decreasing in utility over time. Much as a potential Pareto-improvement in the Kaldor–Hicks sense can yield Pareto optimality when combined with appropriate compensation of losers by winners, so too can dynamic efficiency lead to the ambitious goal of sustainability when combined with appropriate intergenerational transfers. The implication is that much as practical economic analyses often resort to seeking potential Pareto-improvements (see the following section), so too might intertemporal economic analyses focus on dynamic efficiency, leading to the possibility, at least, of sustainability.

---

\(^{10}\) Executive Order 12898 (1994) provides a mandate for Federal agencies to make “environmental justice” part of their missions by considering possible negative effects of proposed policies on minority and low-income populations (U.S. Council on Environmental Quality, 1997). In the international realm, as early as 1987, the Brundtland Commission defined development as sustainable “when it meets the needs of the present without compromising the ability of future generations to meet theirs” (World Commission on Environment and Development, 1987).

\(^{11}\) In the intertemporal realm, Rawls’ (1971) cognitive device of a “veil of ignorance” is insightful, but not operational. Farrow (1998) proposed a modified benefit–cost test for intergenerational equity that emphasized actual compensation rather than potential improvement.

\(^{12}\) For example, Arrow et al. (2004) make a clear distinction between “optimality,” defined as the maximized discounted present value of future well being, and sustainability, defined as “the maintenance or improvement of well being over time.” One exception is provided by Asheim, Buchholz, and Tungodden (2001), who impose so-called efficiency and equity axioms and show that if social preferences fulfill these two axioms, any optimal path will lead to an efficient and non-decreasing path, thus implicitly including dynamic efficiency in the definition of sustainability. For a broader discussion of sustainability and optimality, see Pezzey (1992) and Weitzman (2003), and for a review of the major issues involved, see Pezzey and Toman (2001, 2002).
2.1.2. Benefit–cost analysis of environmental regulations

While conceptually straightforward, the soundness of empirical benefit–cost analysis rests upon the availability of reliable estimates of social benefits and costs, including estimates of the social discount rate.

2.1.2.1. Discounting

Decisions made today typically have impacts both now and in the future. In the environmental realm, many of the future impacts are from policy-induced improvements, and so in this context, future benefits (as well as costs) of policies are discounted (Goulder and Stavins, 2002). The present value of net benefits (PVNB) is defined as:

\[ PVNB = \sum_{t=0}^{T} \left\{ (B_t - C_t) \cdot (1 + r)^{-t} \right\} \]

where \( B_t \) are benefits at time \( t \), \( C_t \) are costs at time \( t \), \( r \) is the discount rate, and \( T \) is the terminal year of the analysis. A positive PVNB means that the policy or project has the potential to yield a Pareto improvement (meets the Kaldor–Hicks criterion). Thus, carrying out benefit–cost or “net present value” (NPV) analysis requires discounting to translate future impacts into equivalent values that can be compared. In essence, the Kaldor–Hicks criterion provides the rationale both for benefit–cost analysis and for discounting.

Choosing the discount rate to be employed in an analysis can be difficult, particularly where impacts are spread across a large number of years involving more than a single generation. The rate chosen can have a significant effect if there are large differences among policies in the timing of benefits and costs. In general, benefits and costs should be discounted at the social discount rate—the relative valuation placed by society on future consumption presently sacrificed. In theory, the social discount rate could be derived by aggregating the individual time preference rates of all parties affected by a policy. Under idealized conditions, the market interest rate would reflect the benefits.

13 Early volumes on benefit–cost analysis include those by Mishan (1976) and Stokey and Zeckhauser (1978); and a recent text is by Boardman et al. (2001). One of the earliest applications to environmental and natural resource policy was by Eckstein (1958).

14 Neither benefit/cost ratios (dividing benefits by costs) nor internal rates of return (the interest rate that results in the present value of benefits being equal to the present value of costs) provide satisfactory alternatives to the net present value criterion, because—among other reasons—neither takes into account scale, and hence both can fail to make proper comparisons among policies using the Kaldor–Hicks criterion. Benefit–cost ratios have the additional problem that the ranking of projects is sensitive to the fundamentally arbitrary judgment of whether an environmental externality is considered to be an increment to costs or a decrement to benefits.

15 Useful surveys include Lind (1982) and Portney and Weyant (1999). An important distinction is whether a publically-mandated policy or project calls for public or private spending. On the effects of this distinction on the choice of discount rate, see, for example: Scheraga and Sussman (1998).
marginal rate of time preference of individuals, but the presence of taxes, risk, liquidity constraints, limited information, and other imperfections means that the social discount rate is not reflected by any particular market rate (Newell and Pizer, 2004).

Alternatives to constant exponential discounting have received consideration. Evidence from market behavior and from experimental economics indicates that individuals may employ lower discount rates for impacts of larger magnitude, higher discount rates for gains than for losses, and rates that decline with the time span being considered (Cropper, Aydede, and Portney, 1994; Cropper and Laibson, 1999). In particular, there has been both empirical and theoretical support for the use of hyperbolic discounting and similar approaches with declining discount rates over time (Ainslie, 1991; Weitzman, 1994, 1998), but most of these approaches suffer from the problem that they would imply inconsistent decisions over time. Declining discount rates based on uncertainty in future rates, however, need not suffer from the time-inconsistency problem (Newell and Pizer, 2003a).

The choice of discount rate can be particularly important in the case of environmental problems with very long time horizons, such as global climate change, radioactive waste disposal, groundwater pollution, and biodiversity preservation (Revesz, 1999). Choosing an intergenerational rate is difficult, because the preferences of future generations are unknowable, and ethical questions arise about trading off the well-being of future generations. Approaches to intergenerational discounting have been considered in two conceptual categories. One relies on a social planner approach, which seeks to maximize the utilities of present and future generations, based on a social welfare function (Lind, 1995; Schelling, 1995; Arrow et al., 1996a). A second approach is based on the preferences of existing individuals, and assumes that one of the allocation decisions these individuals must make is about the welfare of future generations (Shefrin and Thaler, 1988; Cropper, Aydede, and Portney, 1992; Rothenberg, 1993; Schelling, 1995).

What discount rates are actually employed by government agencies? The general answer is a “large range.” For many years, the U.S. Office of Management and Budget (OMB) required the use of a 7 percent real rate for Regulatory Impact Analyses (RIAs), despite the fact that this seems high compared with advice from economists regarding the social discount rate, which would place it in the range of 2 to 3 percent. Why did this persist? One possible rationale was that OMB believed that agencies want their policies, programs, and projects to go forward, and so will tend to exaggerate benefits relative to costs, and that OMB tried to counteract this effect by using a higher discount rate. In any event, reforms put in place by OMB in September of 2003 included the use of a 3 percent real discount rate for intragenerational analyses and lower, unspecified rates for intergenerational contexts (U.S. Office of Management and Budget, 2003).

Several general principles are worth noting. First, it is generally appropriate to employ the same discount rate for benefits and costs. Second, if private capital investments

16 This rationale assumes that the policies in question have the time profile of typical investments, that is, up-front costs and delayed benefits.
will be displaced by public projects, this should be taken into account in estimates of future benefits and costs prior to discounting. Third, estimates of future benefits and costs that may be uncertain or involve risk should be adjusted accordingly (such as through the use of certainty-equivalents), but the discount rate itself should not be changed to account for risk or uncertainty. Fourth, sensitivity analysis using alternative discount rates should be carried out.

2.1.2.2. Benefit concepts and taxonomies

If an environmental change matters to any person—now or in the future—then it should, in principle, show up in an economic assessment.\(^{17}\) Thus, the economic concept of environmental benefits is considerably broader than most non-economists would think.\(^{18}\) The environment can be viewed as a form of natural asset that provides service flows used by people in the production of goods and services, such as agricultural output, human health, recreation, and more amorphous goods such as quality of life. This effect is analogous to the manner in which real physical capital assets provide service flows used in manufacturing. As with real physical capital, a deterioration in the natural environment (as a productive asset) reduces the flow of services the environment is capable of providing.

Protecting the environment usually involves active employment of capital, labor, and other scarce resources. Using these resources to protect the environment means they are not available to be used for other purposes. Hence, the economic concept of the value or benefit of environmental goods and services is couched in terms of society’s willingness to make trade-offs between competing uses of limited resources, and in terms of aggregating over individuals’ willingness to make these trade-offs. Thus, the benefits of an environmental policy are defined as the collection of individuals’ willingness to pay (WTP) for the reduction or prevention of environmental damages or individuals’ willingness to accept (willingness to accept (WTA) compensation to tolerate such environmental damages. In theory, which measure of value is appropriate for assessing a particular policy depends upon the related assignment of property rights, the nature of the status quo, and whether the change being measured is a gain or a loss, but under a variety of conditions, the difference between the two measures may be expected to be relatively small (Willig, 1976).\(^{19}\) Empirical evidence suggests larger than expected differences between willingness to pay and willingness to accept (Cummings, Brookshire, and Schulze, 1986; Horowitz and McConnell, 2002, 2003).

---

\(^{17}\) This reflects the anthropocentric view employed in economics, which does not include welfare incurred by other living creatures, unless it (indirectly) affects humans. For a recent argument involving non-anthropocentric values, see Ariansen (1998).

\(^{18}\) For a summary of myths that non-economists seem to have regarding economics in the environmental realm, and a set of responses thereto, see: Fullerton and Stavins (1998).

\(^{19}\) The difference depends on the magnitude of the impact, as well as the related income elasticity of demand. Consumers’ surplus, derived from the observed Marshallian demand curve, provides a close approximation for equivalent and compensating variations (Willig, 1976). Willig’s analysis is of price changes, but Randall and Stoll (1980) showed that similar results hold for quantity changes. For reviews of empirical studies of willingness-to-pay and willingness-to-accept measures, see: Horowitz and McConnell (2002, 2003). For examinations of the relationship between Willig’s conditions and weak complementarity, see: Bockstael and McConnell (1993); Palmquist (2005); and Smith and Banzhaf (2004).
Fisher, McClelland, and Schulze, 1988). Theoretical explanations include psychological aversion to loss and poor substitutes for environmental amenities. In particular, Hanneman (1991) demonstrated that for quantity changes, the less perfect the substitutes that are available for a public good, the greater the expected disparity between WTP and WTA.

The benefits people derive from environmental protection are numerous and diverse. From a biophysical perspective, such benefits can be categorized as being related to human health (mortality and morbidity), ecological impacts (both market and non-market), or materials damage. From an economic perspective, a critical distinction is between use value and non-use value. In addition to the direct benefits (use value) people receive through protection of their health or through use of a natural resource, people also derive passive or non-use value from environmental quality, particularly in the ecological domain. For example, an individual may value a change in an environmental good because she wants to preserve the option of consuming it in the future (option value) or because she desires to preserve the good for her heirs (bequest value). Still other people may envision no current or future use by themselves or their heirs, but still wish to protect the good because they believe it should be protected or because they derive satisfaction from simply knowing it exists (existence value).

2.1.2.3. Cost concepts and taxonomies In the environment context, the economist’s notion of cost, or more precisely, opportunity cost, is a measure of the value of whatever must be sacrificed to prevent or reduce the risk of an environmental impact. Hence, the costs of environmental policies are the forgone social benefits due to employing scarce resources for environmental policy purposes, instead of putting those resources to their next best use.

A taxonomy of environmental costs can be developed, beginning with the most obvious and moving towards the least direct (Jaffe et al., 1995). First, many policy makers and much of the general public identify the on-budget costs to government of administering (monitoring and enforcing) environmental laws and regulations as the cost of environmental regulation. This meets the notion of opportunity cost, since administering environmental rules involves the employment of resources (labor and capital) that could

---

20 It is important to distinguish between ecosystem functions (for example, photosynthesis) and the environmental services produced by ecosystems that are valued by humans (Freeman, 1997). The range of these services is great, including obvious environmental products such as food and fiber, and services such as flood protection, but also including the quality of recreational experiences, the aesthetics of the landscape, and such desires (for whatever reasons) as the protection of marine mammals.

21 Option value and existence value should not be thought of as being additive, since option value is defined from a framework that holds expected utility constant; this is not the case with existence (and bequest) value. The contemporary concept of non-use value relates to what was previously most often characterized as existence value. See: Graham (1981); Bishop (1982); and Smith (1987).

22 Costs and benefits are thus two sides of the same coin. The cost of an environmental-protection measure may be defined as the gross decrease in benefits (consumer and producer surpluses) associated with the measure and with any price and/or income changes that may result (Cropper and Oates, 1992).
otherwise be used elsewhere. But economic analysts also include as costs the capital and operating expenditures associated with regulatory compliance. Indeed, these typically represent a substantial portion of the overall costs of regulation, although a considerable share of compliance costs for some regulations falls on governments rather than private firms. Additional direct costs include legal and other transaction costs, the effects of refocused management attention, and the possibility of disrupted production.

Next, there are what have sometimes been called “negative costs” of environmental regulation, including the beneficial productivity impacts of a cleaner environment and the potential innovation-stimulating effects of regulation. General equilibrium or multi-market effects associated with discouraged investment and retarded innovation constitute another important layer of costs, as do the transition costs of real-world economies responding over time to regulatory changes.

2.1.2.4. Cost estimation methods The merits of alternative empirical cost estimation methods are related to the magnitude of the various categories of costs outlined above. Methods of direct compliance cost estimation, which measure the costs to firms of purchasing and maintaining pollution-abatement equipment plus costs to government of administering a policy, are acceptable when behavioral responses, transitional costs, and indirect costs are small. Partial and general equilibrium analysis allow for the incorporation of behavioral responses to changes in public policy. Partial equilibrium analysis of compliance costs incorporates behavioral responses by modeling supply and/or demand in major affected markets, but assumes that the effects of a regulation are confined to one or a few markets. This may be satisfactory if the markets affected by the policy are small in relation to the overall economy, but if an environmental policy is expected to have large consequences for the economy, general equilibrium analysis is required.

General equilibrium cost estimation methods include both input-output models and computable general equilibrium models. Input-output analysis quantifies the flow of goods and services in an economy using fixed-coefficient relationships (Leontief, 1966, 1970), and is limited in its usefulness by restrictive assumptions of constant returns to scale, fixed prices, and fixed producer and consumer behavior. Computable general equilibrium (CGE) models relax these assumptions at the cost of greater data

---

23 One example in the United States is the federal regulation of contaminants in drinking water, the cost of which is borne primarily by municipal governments.

24 The notion that environmental regulation can foster economic growth is a controversial one among economists. For a debate on this proposition, see: Porter and van der Linde (1995); and Palmer, Oates, and Portney (1995).

25 For example, if a firm chooses to close a plant because of a new regulation (rather than installing expensive control equipment), this would be counted as zero cost in narrow compliance-cost estimates, but it is obviously a real cost.

26 If a policy will result in only small changes in consumer and producer behavior, real resource and regulatory costs will represent the bulk of costs. But when behavioral responses are expected to be sizeable, social welfare costs associated with losses in consumer and producer surplus due to a rise in prices or a decrease in output can be significant.
requirements. The potential importance of tax-interaction effects (Goulder, Perry, and Burtraw, 1997), described below in section 3.1.3.3, highlight the value of employing CGE models for comprehensive cost analysis.

How well have cost estimation methods performed in practice? In a retrospective examination of 28 environmental and occupational safety regulations, Harrington, Morgenstern, and Nelson (2000) found that fourteen had produced ex ante cost estimates that exceeded actual ex post costs. But these errors were mainly due to overestimates of the quantity of emissions reduction that would occur. In terms of per-unit abatement costs, overestimation and underestimation were equally common, although for regulations that used economic incentives, per-unit costs were consistently overestimated. Harrington, Morgenstern, and Nelson (2000) attributed this to technological innovation which was stimulated by these market-based instruments, and which thereby reduced abatement costs.

2.1.2.5. Benefit estimation methods Empirical methods of economic valuation were originally developed in the context of changes in individuals’ incomes and in prices faced in the market. Over the past thirty years, these methods have been extended to accommodate changes in the qualities of goods, to public goods that are shared by individuals, and to other non-market services such as environmental quality and human health. With markets, consumers’ decisions about how much of a good to purchase at different prices reveal useful information regarding the surplus consumers gain. With non-market environmental goods, it is necessary to infer this willingness to trade off other goods (or monetary amounts) for environmental services using other methods. A repertoire of techniques has been developed in two broad categories: revealed preference (indirect measurement) and stated preference (direct questioning).

2.1.2.5.1. Revealed preference methods Whenever possible, it is preferable to measure trade-offs by observing actual decisions made by consumers in real markets. In limited situations, researchers can observe relationships that exist between a non-marketed, environmental good and some good that has a market price. In this case, individuals’ decisions to avert or mitigate the consequences of environmental deterioration can shed light on how people value environmental quality (averting behavior estimates). In other cases, individuals reveal their preferences for environmental goods in the housing market (hedonic property value methods), or for related health risks in labor markets (hedonic wage methods). In still other cases, individuals reveal their demand for recreational amenities through their decisions to travel to specific locations.

27 For a recent survey of computable general equilibrium models, see Conrad (2002), and for an application of CGE modeling to estimate the costs of the Clean Air and Clean Water Acts, see Hazilla and Kopp (1990).
28 Three of the ex ante cost analysis were underestimates; the other eleven were approximately correct.
29 On this, also see: Hammitt (2000).
30 For an intellectual history of developments in this area, see Cropper (2000), and for a survey of theoretical underpinnings and empirical issues associated with alternative benefit estimation methods, see Freeman (2003).
(Hotelling–Clawson–Knetsch and related recreation-demand methods). In addition, empirical evidence of environmental benefits may be obtained when individuals express their willingness to pay for a privately-traded option to use a freely-available public good. This set of revealed preference methods can be used to estimate the trade-offs that are at the heart of environmental valuation, but—as explained below—the scope of potential application of these methods is limited.

The averting behavior method, in which values of willingness to pay are inferred from observations of people’s behavioral responses to changes in environmental quality, is grounded in the household production function framework. People sometimes take actions to reduce the risk (averting behavior) or lessen the impacts (mitigating behavior) of environmental damages, for example, by purchasing water filters or bottled water. In theory, people’s perceptions of the cost of averting behavior and its effectiveness should be measured (Cropper and Freeman, 1991), but in practice actual expenditures on averting and mitigating behaviors are typically employed, with the results sometimes interpreted as constituting a lower bound on willingness to pay. Such an interpretation, however, can be misleading (Shogren and Crocker, 1991, 1999). An additional challenge is posed by the necessity of disentangling attributes of the market good or service. For example, bottled water may be safer, taste better, and be more convenient. In this case, willingness to pay for safer water might be overestimated by an averting behavior approach. On the other hand, since bicycle helmets are uncomfortable, expenditures on such equipment could lead to an underestimate of willingness to pay for risk reduction.

Hedonic pricing methods are founded on the proposition that people value goods in terms of the bundles of attributes that constitute those goods. In theory, the value of the environmental component of a particular good can be extracted by statistically decomposing the value of the total good into willingness to pay for multiple attributes. In the environmental sphere, hedonic methods have been applied to property values and to wages.

Hedonic property value methods employ data on residential property values and home characteristics, including structural, neighborhood, and environmental quality attributes. By regressing the property value on key attributes, the hedonic price function

---

31 See Becker (1965) for the early development of the household production method, and Bockstael and McConnell (1983) for the conditions under which the benefits of a public good can be inferred from the demand function for a related private, market good. Mäler’s (1985) theoretical explication builds upon a proposal by Ridker (1967). For an early application to human health, see Grossman (1972), and for a complete theoretical exposition, see Courant and Porter (1981).

32 The hedonic pricing method was originated by Waugh (1928), but it was Court (1939) who developed the method using multiple regression techniques. Hedonic pricing was revived and further developed econometrically by Griliches (1961). Most applications in the environmental realm stem from Rosen (1974). For an examination of the conditions under which the results from the hedonic price function can be used for benefit estimation, see: Bartik (1988a). More recent treatments include those by Ekeland, Heckman, and Nesheim (2002, 2004), which reflect on the identification issues originally addressed by Brown and Rosen (1982). Surveys are provided by Palmquist (1991) and Taylor (2003).

33 For surveys of methodological developments and applications of hedonic property methods, see: Bartik and Smith (1987); and Palmquist (2005).
is estimated:

\[ P = f(\bar{x}, \bar{z}, e) \]  

(3)

where \( P \) = housing price (includes land);
\( \bar{x} \) = vector of structural attributes;
\( \bar{z} \) = vector of neighborhood attributes; and
\( e \) = environmental attribute of concern.

From the estimated hedonic price function of equation (3), the marginal implicit price of any attribute, including environmental quality, can be calculated as the partial derivative of the housing price with respect to the given attribute:

\[ \frac{\partial P}{\partial e} = \frac{\partial f(\cdot)}{\partial e} = P_e \]  

(4)

This marginal implicit price, \( P_e \), measures the aggregate marginal willingness to pay for the attribute in question, and it may be interesting in and of itself. For purposes of benefit estimation, however, the demand function for the attribute is required, and so it becomes necessary to examine how the marginal implicit price of the environmental attribute calculated from equation (4) varies with changes in the quantity of the attribute and other relevant variables. If the hedonic price equation (3) is non-linear, then fitted values of \( P_e \) can be calculated as \( e \) is varied, and a second-stage equation can be estimated:

\[ \hat{P}_e = g(e, \bar{y}) \]  

(5)

where \( \hat{P}_e \) = the fitted value of the marginal implicit price of \( e \) from the first-stage equation; and
\( \bar{y} \) = a vector of factors that affect marginal willingness to pay for \( e \), including buyer characteristics.

Equation (5), above, has been interpreted as the demand function for the environmental attribute—from which benefits (consumers surplus) can be estimated in the usual way—but there are several important issues and problems.

Most important among the problems confronting the use of the hedonic property method for environmental benefit estimation is the question of whether a demand function has actually been estimated, since environmental quality may affect both the demand for housing and the supply of housing. Thus, the classic identification problem of econometrics arises. In addition, hedonic property value methods build upon a model of the housing market that is in equilibrium for all attributes, with buyers and sellers having full information. Informational asymmetries may distort the analysis, particularly if perceptions about environmental attributes are different from scientific measurements of these values. That is, if individuals’ perceptions of environmental attributes do not correspond to actual measurement of attributes, then estimated marginal implicit prices will be biased (possibly upward, possibly downward, depending upon the nature of perceptions).
Because the hedonic property method is based on analysis of marginal changes, it should not be applied to analysis of policies with large anticipated effects, and because the method’s data requirements are considerable, omitted variable bias may be a problem. Finally, although the method seems very well suited for some environmental attributes, including noise abatement and proximity to waste sites, many other environmental amenities do not lend themselves to this type of analysis.

A related benefit-estimation method frequently employed in the environmental policy domain is the hedonic wage method, which is based on the empirical reality that individuals in well functioning labor markets make trade-offs between wages and risk of on-the-job injuries (or death). In a hedonic context, a job is a bundle of characteristics, including its wage, responsibilities, and risk, among others factors.34 Two jobs that require the same skill level but have different risks of on-the-job mortality will pay different wages. On the labor supply side, employees tend to require extra compensation to accept jobs with greater risks; and on the labor demand side, employers are willing to offer higher wages to attract workers to riskier jobs. Hence, labor market data on wages and job characteristics can be used to estimate econometrically people’s marginal implicit price of risk, that is, their valuation of risk. By regressing the wage on key attributes, the hedonic price function is estimated:

\[ W = h(\chi, r) \]  

(6)

where \( W \) = wage (in annual terms); 
\( \chi \) = vector of worker and job characteristics; and 
\( r \) = mortality risk of job.

The marginal implicit price of risk is calculated as the partial derivative of the annual wage with respect to the measured mortality risk:

\[ \frac{\partial W}{\partial r} = \frac{\partial h(\cdot)}{\partial r} = W_r \]  

(7)

Note that the marginal implicit price of risk is the average annual income necessary to compensate a worker for a marginal change in risk throughout the year. This marginal implicit price varies with the level of risk.

Many of the issues that arise with the hedonic property value method have parallels here. First, there is the possibility of simultaneity: causality between risk and wages can run in both directions. For example, higher ambient air pollution might lead to higher compensating wages, but higher wages and incomes in an area may lead to more automobiles and hence more air pollution. Also, if individuals’ perceptions of risk do not correspond with actual risks, then the marginal implicit price of risk calculated from a hedonic wage study will be biased, although, as before, the direction of the bias is not obvious. Imperfections in labor markets (less than perfect mobility) can cause problems,

34 For a detailed treatment of the hedonic wage model, see Viscusi (1992, 1993).
but more important are the significant data requirements that can lead to omitted variable bias.35

Direct applications of the method in the environmental realm would appear to be severely limited. Indeed, direct application is limited to occupational, as opposed to environmental exposures and risks. Yet hedonic wage methods are of considerable importance in the environmental policy realm, because the results from hedonic wage studies have frequently been used through “benefit transfer” to infer the value of a statistical life (VSL), as we discuss below in section 2.1.2.5.5.

Recreational activities represent a potentially large class of benefits that are particularly important in assessing policies affecting the use of public lands. The models used to estimate recreation demand fall within the class of household production models, discussed above. First, travel cost models (or Hotelling–Clawson–Knetsch models) use information about time and money spent visiting a site to infer the value of that recreational resource. The simplest version of the method involves one site and uses data from surveys of users from various geographic origins, together with estimates of the cost of travel and opportunity cost of time to infer a demand function relating the number of trips to the site as a function of people’s willingness to pay for the experience.36

The most significant limitation of the simplest travel-cost model is the omission of potential substitute sites. Although one obvious approach is to include the price (travel and opportunity cost) of substitute sites as additional independent variables, better approaches involve multi-site travel cost models or the use of random utility models. Random utility models explicitly model the consumer’s decision to choose a particular site from alternative recreation locations, assessing the probability of visiting each location. The most important attribute of random utility models is that they can be used to value changes in environmental quality by comparing decisions to visit alternative sites.37

All recreation demand models share a set of limitations. First, the valuation of costs depends critically on empirical estimates of the opportunity cost of (leisure) time, which is notoriously difficult to estimate. Also, most trips to a recreation site are part of a multi-purpose experience. If this is ignored, willingness to pay will be over-estimated. In addition, random utility models rely on people’s perceptions of environmental quality changes, and so changes that are difficult to observe may be valued “incorrectly.”

35 If individuals change jobs and homes simultaneously—not an unreasonable expectation in some cases—then the observed marginal willingness to pay will reflect both the labor and property markets. On this, see: Rosen (1979); Roback (1982); and Bartik and Smith (1987).

36 The conceptual approach was proposed by Harold Hotelling in a 1954 letter to the Director of the U.S. National Park Service, and the method was subsequently developed and applied by Davis (1963) and Clawson and Knetsch (1966). For a survey of travel cost models, see Bockstael (1996); and for a recent survey of recreation demand models, see Phaneuf and Smith (2005).

37 For detailed treatments of random utility/discrete choice models, see: Bockstael, Hanemann, and Strand (1986); Herriges and Kling (1999); Haab and McConnell (2002); and Parsons (2003).
Finally, like all revealed-preference approaches, recreation demand models can be used to estimate use value only; non-use value cannot be examined.\textsuperscript{38} An alternative approach to assessing people’s willingness to pay for recreational experiences is to draw on evidence from private options to use public goods. This approach also fits within the household production framework, and is based upon the notion of estimating the derived demand for a privately traded option to utilize a freely-available public good. In particular, the demand for state fishing licenses has been used to infer the benefits of recreational fishing (\textit{Snyder, Stavins, and Wagner, 2003}). Using panel data on state fishing license sales and prices for the continental United States over a fifteen-year period, combined with data on substitute prices and demographic variables, a license demand function was estimated, from which the expected benefits of a recreational fishing day were derived.

In summary, revealed-preference methods of environmental benefit estimation are based upon sound theoretical foundations and can be empirically effective. If well executed, these methods can produce relatively accurate (that is, unbiased) and relatively precise (that is, low variance) estimates. These approaches are therefore strongly favored by economists, both on theoretical and empirical grounds. But revealed preference methods are severely limited in the scope of their direct applicability. In many situations, it is simply not possible to observe behavior that reveals people’s valuations of changes in environmental goods and services. This is particularly true with non-use values. With no standard market trade-offs to observe, economists must resort to surveys in which they construct hypothetical markets, employing stated preference, as opposed to revealed preference methods.

2.1.2.5.2. Stated preference methods  In the best known stated preference method, \textit{contingent valuation}, survey respondents are presented with scenarios that require them to trade-off, hypothetically, something for a change in the environmental good or service in question. Stated preference methods depend on directly questioning individuals about their valuation of changes in environmental quality. While controversial because of the potential for biased answers, based on intentions rather than actions, stated preference methods are the only way to estimate non-use values for environmental goods.

Contingent valuation (CV) presents survey subjects with a hypothetical increase or decrease in environmental quality and asks how much they would be willing to pay or accept to enact or prevent such changes. The essential steps in carrying out a CV study are: clearly defining the good or service and the policy-induced change in the good or service to be valued; identifying the geographical scope of the “market;” conducting focus groups on components of the survey; pretesting the survey instrument; administering the survey to a random sample of the market; testing the results for reliability (bias) and validity (theoretical correspondence); and possibly using the elicited

\textsuperscript{38} On the possibility of using corner-solution models of recreational behavior to estimate non-use values (employing important assumptions along the way), see: \textit{Herriges, Kling, and Phaneuf (2004)}. 
willingness-to-pay data for various quantities of the good/service to construct a demand function, and estimate benefits.\textsuperscript{39}

The CV survey instrument itself is used to: collect information on the consumer’s past, present, and expected future use of the environmental good (or service); collect information on the respondent’s socioeconomic characteristics; present a hypothetical scenario describing a change in the good to be valued; present a specific hypothetical payment vehicle, which is both plausible and understandable (examples include taxes, user fees, and product prices); and elicit the respondent’s willingness to pay, reminding the respondent of the existence of substitutes.

Elicitation methods have been of four principal types. First, the simplest approach is to ask people for their maximum willingness to pay, but there are few real markets in which individuals are actually asked to generate their reservation prices, and so this method is considered unreliable. Second, in a bidding game, the researcher begins by stating a willingness-to-pay number, asks for a yes-no response, and then increases or decreases the amount until indifference is achieved. The problem with this approach is the inevitable introduction of significant starting-point bias. Third, a related approach is the use of a payment card to be shown to the respondent, but the problem here is that the range of WTP on the card may still introduce bias, and the approach cannot be used with telephone surveys. Fourth and finally, the referendum (discrete choice) approach is favored by researchers. Here, each respondent is offered a different WTP number, to which a simple yes-no response is solicited. This approach minimizes bias, but requires considerably more observations.

The primary advantage of contingent valuation is that it can be applied to a wide range of situations, including use as well as non-use value, but potential problems remain. First, respondents may not understand what they are being asked to value. This may introduce greater variance, if not bias in responses. Likewise, respondents may not take the hypothetical market seriously, because no budget constraint is actually imposed. This can increase variance and bias. On the other hand, if the scenario is “too realistic,” strategic bias may be expected to show up in responses. Finally, the “warm glow effect” may plague some stated preference surveys: people may purchase moral satisfaction with large, but unreal statements of their willingness-to-pay (Andreoni, 1995). For example, in one CV study, it was found that 63 percent of respondents indicated they were willing to pay $30 to a specific leading Norwegian environmental organization to protect resources. But when the same organization followed up with mail solicitations to the same sample, fewer than 10 percent of the original respondents contributed anything (Seip and Strand, 1992).

The 1989 Exxon Valdez oil spill in Prince William Sound off the coast of Alaska led to massive litigation, and—as a consequence—resulted in the most prominent use

\textsuperscript{39} For a comprehensive treatment of contingent valuation methods, see Mitchell and Carson (1989). For more recent surveys, see: Brown (2003); and Boyle (2003).
ever of the concept of non-use value and the method of contingent valuation for its estimation. The result was a symposium sponsored by Exxon attacking the CV method (Hausman, 1993), and the creation of a government panel—established by the National Oceanic and Atmospheric Administration (NOAA) and chaired by two Nobel laureates in economics—to assess the scientific validity of the CV method (Arrow et al., 1993). The NOAA panel concluded that “CV studies can produce estimates reliable enough to be the starting point of a judicial process of damage assessment, including lost passive (non-use) values” (Arrow et al., 1993, p. 4610). The panel offered its approval of CV methods subject to a set of best-practice guidelines. Since that time, economists have continued to seek ways to improve CV methods and to verify reliability through: replication of CV results; comparison of CV results with other estimates (Hanneman, 1994); and—where possible—comparison of CV results with actual behavior. Nevertheless, some economists remain highly skeptical of this method.

2.1.2.5.3. Fallacious methods of “benefit estimation” It is important to distinguish the averting behavior method, described above, from so-called “avoided-cost measures of benefits” in general, which are attempts to substitute for real measures of benefits the cost of the next most costly means of achieving some goal. Unless individuals have demonstrated their willingness to undertake voluntarily the alternative activities—as in the case of averting behavior methods—using costs as proxies for benefits is illegitimate; it simply converts what would be a benefit–cost comparison into a cost-cost (that is, cost-effectiveness) comparison. By applying “avoided-cost measures of benefits,” any proposed project can be made to appear desirable. By taking the next most costly approach of achieving an objective and calling that the project’s benefits, one will always find that “benefits”—so measured—exceed costs.

Related to attempts to substitute costs for true measure of benefits is the so-called “societal revealed preference” (SRP) approach, whereby analysts seek to infer the benefits of a proposed policy from the costs of previous regulatory actions. Of course, true revealed preference benefit estimation methods require that individuals or groups voluntarily undertake actions and pay the costs of undertaking those actions. The SRP method fails this test. Only if the previous regulation itself passed a benefit–cost test could the costs of that regulation possibly be assumed to have any particular relation to its benefits. The SRP method is not a revealed-preference method, and indeed is not a benefit-estimation method at all, but—at most—a cost-effectiveness comparison. The purpose of a benefit–cost analysis is to assess policies by contrasting their benefits and their costs; the SRP approach reverses this, taking the fact that a policy exists as evidence that its benefits exceed its costs (and therefore that its benefits can be proxied by

---

40 The CV study carried out to estimate non-use value lost as a result of the Exxon Valdez oil spill was eventually published (Carson et al., 2003).

41 See Portney (1994) for an overview of the debate, Diamond and Hausman (1994) for a skeptical view, and Hanneman (1994) for a defense of CV methods. More recent contributions include: Carson, Flores, and Hanemann (1998); Cameron et al. (2002); and Champ, Boyle, and Brown (2003).
its costs, at a minimum). The use of such approaches would stand the very process of regulatory impact analysis on its head.

Finally, an approach frequently used by government agencies and others in attempts to value changes in morbidity (non-fatal health effects) is the so-called cost of illness method (U.S. Environmental Protection Agency, 2002). This approach does not provide a theoretically correct measure of willingness to pay or willingness to accept, but instead measures explicit market costs resulting from changes in the incidence of illness. Direct and indirect medical costs are included, where direct costs refer to diagnosis, treatment, and rehabilitation, and indirect costs refer to the loss of productivity attributable to illness. “Pain and suffering” and averting behavior are not included. Cost-of-illness estimates have therefore been interpreted as providing a lower bound on willingness to pay (Harrington and Portney, 1987), but this may not be the case because the reality of individuals passing costs on to third parties (insurers, hospitals, and employers) means that costs incurred may overstate true individual willingness to pay.

2.1.2.5.4. Benefit transfer Because of the considerable time and cost of both revealed-preference and stated-preference valuation methods, government agencies—including the U.S. Environmental Protection Agency—frequently rely on the transfer of existing estimates from previous research (the “study case”) to new contexts (the “policy case”). Such benefit-transfers are very inexpensive compared with original research, but the estimates are far less accurate and reliable, and inevitably introduce arbitrary elements of judgment into the analysis.

Three principal benefit transfer methods have been employed. First, point estimates involve the simple adoption of a benefit number from a previous study. This approach is generally considered unacceptable. Second, a benefit function may be adopted from the study case, plus values of exogenous variables from the policy case; then benefits can be estimated. Third, if such benefit functions are not available from previous research, a meta-analysis may be carried out, combining values from a variety of previous studies, estimating a statistical relationship of the factors affecting benefits, and then employing values of exogenous variable from the policy case in order to estimate (the fitted value of) benefits (U.S. Environmental Protection Agency, 1993, 2000a).

Two major criteria are useful for judging benefit-transfer exercises (Desvousges, Johnson, and Banzhaf, 1998). The first criterion is soundness—was the study-case analysis itself of sufficient quality? The second criterion is similarity. The basic commodities analyzed in the study case and the policy case should be essentially equivalent; the baselines and the degrees of change in the environmental good or service should be similar; and the affected populations should be similar. This is particularly challenging in the natural resources context, because values tend to be highly dependent upon location, suggesting the infeasibility of meeting the similarity condition (Rosenberger and Loomis, 2003).

2.1.2.5.5. Valuing mortality risk reductions How much would individuals sacrifice to achieve a small reduction in the probability of death during a given period of time? How much compensation would individuals require to accept a small increase in that
probability? These are reasonable economic questions, given the fact that most environmental regulatory programs result in small changes in individuals’ mortality risks. Empirical methods, considered above, including hedonic wages studies, averted behavior, and contingent valuation, can provide estimates of marginal willingness to pay or willingness to accept for small changes in mortality risk. For purposes of benefit transfer, such estimates have been normalized into measures of the “value of a statistical life” (VSL).\(^42\)

The VSL is not the value of an individual life—neither in ethical terms, nor in technical, economic terms. Rather it is simply a convention:\(^43\)

\[
VSL = \frac{\text{MWTP or MWTA (from hedonic wage or CV)}}{\text{Small Risk Change}}
\]

where MWTP and MWTA, respectively, refer to marginal willingness to pay and marginal willingness to accept. For example, if people are willing, on average, to pay $12 for a risk reduction from 5 in 500,000 to 4 in 500,000, equation (8) would yield:

\[
VSL = \frac{12}{0.000002} = 6,000,000
\]

Thus, VSL quantifies the aggregate amount that a group of individuals are willing to pay for small reductions in risk, standardized (extrapolated) for a risk change of 1.0.\(^44\)

It is not the economic value of an individual life. The VSL calculation above does not signify that an individual would pay $6 million to avoid (certain) death this year, or accept (certain) death this year in exchange for $6 million. It does imply that 100,000 similar people would together pay $6 million to eliminate the risk that is expected to kill one of them randomly this year.\(^45\)

There has been considerable debate regarding whether and how VSLs should be adjusted for risk characteristics, including the latency periods of pre-mortality illness, the dread associated with some forms of mortality, and the difference between voluntary and involuntary risk. Discounting has been the usual way of handling any latency period prior to mortality, but this may oversimplify how individuals value future impacts.

\(^42\) For comprehensive surveys of the VSL literature, see: Fisher, Chestnut, and Violette (1989), Miller (1989), and Viscusi (1993).

\(^43\) The “convention” is to express the marginal willingness to pay for a small reduction in mortality risk or marginal willingness to accept compensation for a small increase in mortality risk, normalized for a risk change of 1.0. It is critical to understand that the convention could just as easily be for a risk change of one in a million. Indeed, if that were the convention, the usefulness of the device for benefit analysis would not be affected in the least, the unfortunate and misleading name of “value of a statistical life” would be avoided, and much of the ensuing controversy might not have arisen. Unfortunately, we are stuck with the normalization and the name, or at least the abbreviation, VSL.

\(^44\) The first formal development of the concept of willingness to pay for mortality risk reductions was by Jones-Lee (1974).

\(^45\) The U.S. Environmental Protection Agency employs a VSL of $6.2 million in Regulatory Impact Analyses. This is the average of 26 (21 hedonic wage and 5 CV) studies upon which EPA draws for its calculation (U.S. Environmental Protection Agency, 2000a).
The dread and pain associated with some forms of mortality is clearly relevant, but is properly considered as a morbidity, not a mortality, effect. Since VSLs draw largely upon hedonic wage studies (see above), they reflect valuations of voluntary risk, but their application to environmental policy assessment is related to involuntary risk.

It is also reasonable to ask whether VSLs should be adjusted for population characteristics. Although there is consistent evidence that mortality risk valuation and income (wealth) are highly correlated, evidence on correlation of valuations and health status is mixed. Perhaps most important, it is expected that people’s willingness to pay for small changes in risk varies over the course of their lives. But the relationship between age and risk valuation is complicated. Standard economic theory would suggest that younger people would have higher values for risk reduction because they have a longer expected life remaining before them and thus a higher expected lifetime utility (Moore and Viscusi, 1988; Cropper and Sussman, 1990). On the other hand, some models and empirical evidence suggest that older people may in fact have a higher demand for reducing mortality risks than younger people, and that the value of a life may follow an “inverted-U” shape over the life-cycle, with its peak during mid-life (Shepard and Zeckhauser, 1982; Jones-Lee, Hammerton, and Philips, 1985; Ehrlich and Chuma, 1990; Krupnick et al., 2002; Mroz and Taylor, 2002; Viscusi and Aldy, 2003; Alberini et al., 2004).

Valuations of non-fatal health effects (morbidity) are also required for many benefit–cost analyses in the environmental realm. The theoretically appropriate measure is aggregate willingness-to-pay to reduce the risk of a given health effect, but—as indicated above—cost-of-illness measurements have been used in administrative and judicial contexts when better estimates were not available. Measuring morbidity effects can be more difficult than estimating mortality impacts because of variations in health endpoints.

2.1.2.6. Critiques of benefit–cost analysis

In addition to criticism (discussed above in section 2.1.1) of the Kaldor–Hicks criterion as a decision rule, there has been considerable criticism of the use of benefit–cost analysis in the environmental realm, both on conceptual and empirical grounds. The most common conceptual objection to benefit–cost analysis from non-economists is that monetary estimates of environmental quality are impossible and/or unethical. Some have argued that the environment has an intrinsic

---


47 See, for example, Desvousges et al. (1996).


value that cannot be quantified numerically, or that attaching a monetary value to environmental quality is ethically wrong because environmental quality should be treated as a basic right that must be protected, regardless of whether the benefits outweigh the costs (Kelman, 1981a). Of course, economic value has a very specific definition: it is a measure of those things that people would be willing to give up to have environmental quality, whether or not it is traded in markets. The implementation of all rights, including those held to be fundamental, requires real resources and imposes real costs. Adding information to the process through benefit–cost analysis can serve to improve decision-making.50

More recently, some critics have questioned the empirical methods used for valuing marginal willingness to pay to avoid and willingness to accept compensation to endure incremental changes in risk of mortality (and morbidity). Unfortunately, some of the most prominent critiques have been premised upon fundamental misunderstanding of those same theories and empirical methods, and have been based upon misleading straw-man caricatures of the positions of economists (Heinzerling, 1998; Ackerman and Heinzerling, 2002).

In this context, although formal benefit–cost analysis should not be viewed as either necessary or sufficient for designing sensible public policy, it can provide an exceptionally useful framework for consistently organizing disparate information, and in this way, it can greatly improve the process and hence the outcome of policy analysis (Arrow et al., 1996b). Economists share concerns about the empirical reliability of benefit–cost estimation methods in specific applications, as highlighted throughout our discussion above. More broadly, economists recognize that while benefit–cost analysis can be very helpful to decision makers, it ought to be considered as an aid to decision makers, not a substitute for decision making.

2.1.3. Other approaches to analyzing the goals of environmental policies

Decision-makers and scholars have proposed other evaluation criteria with which environmental policies might be assessed. One approach, reflected in prevailing interpretations of the Clean Air Act and some other environmental laws, has been to claim to rely solely on biophysical (that is, natural science) information to identify policies that eliminate environmental risks altogether or reduce them to levels deemed acceptable. The Clean Air Act, for example, has been construed to adopt this approach, directing EPA to set its ambient air quality standards at levels that will protect public health with an adequate margin of safety (see section 2.2.2, below). Some legal scholars have defended this view (Heinzerling, 2001), but since many environmental pollutants fail to exhibit clear thresholds below which they pose no health effects, such an approach is unworkable as a normative basis for setting environmental standards. More fundamentally, natural science alone cannot provide a normative basis for setting environmental standards.

50 A brief and pragmatic defense of the use of benefit–cost analysis is provided by Arrow et al. (1996b). Replies to Kelman’s (1981a) critique are provided by DeLong (1981) and Solow (1981).
standards (Coglianese and Marchant, 2004), despite the fact that input from the natural sciences is necessary for implementing economic or most other criteria.

Another problem with a simple risk-elimination approach is that environmental policies can increase certain risks at the same time as they reduce other risks. This motivated the development of risk-risk analysis (Lave, 1981), in which health outcomes of alternative policies are calculated and presented directly, without monetary valuation. An important aspect of the analysis is taking into account both the positive, intended effects on health of the policy under consideration and the negative effects that the policy may bring about. Thus, the analysis compares risk reductions caused by a policy with risks created by the policy. For example, a policy that requires power plants to install pollution abatement equipment may reduce the risk of illness due to environmental pollutants, but increase the risk of on the job injury because of construction needed to meet the standards (Lave, 1981).

Clearly risk-risk analysis cannot be used to ascertain whether a policy fulfills the efficiency criteria, because the only costs counted are other health risks; the real resource costs and opportunity costs of implementing the program are ignored. Furthermore, without a common numeraire, policy-makers have no clear standard for comparing different types of health impacts, and so policies cannot be ranked.51 It has been argued that risk-risk analysis is also flawed because it focuses on negative secondary effects of regulation (ancillary risks), ignoring ancillary benefits (Rascoff and Revesz, 2002). Risk-risk analysis has seen only limited use.

Health-health analysis goes one step further by attempting to quantify resource and opportunity costs, premised on the notion that spending for regulatory programs diverts resources from individuals, causing them to spend less on safety and healthcare, and thereby increasing their morbidity and/or mortality risks.52 Thus, the public health benefits of a program are contrasted with the negative health effects of the program. A common accounting unit is required, typically the number of fatalities. Health-health analysis provides a measure of “net benefits” (lives saved), but this analytical method suffers from a number of severe limitations (Portney and Stavins, 1994): it does not include other benefits besides saved fatalities; the relatively small cost of environmental regulation as a percentage of individual budgets means that there may be no observable effects on individual health expenditures; and accurate analysis depends on the difficult task of estimating the complex empirical relationship between marginal income changes and health risks.

Distributional analysis provides another approach to analyzing the goals of environmental policies in economic terms. Benefit–cost analysis focuses exclusively on aggregate net benefits, and does not take into account the distributional consequences

51 See Viscusi, Magat, and Huber (1991), and Graham and Wiener (1995).
52 Wildavsky (1980) was one of the first to describe the relationship between regulation and increased morbidity or mortality due to loss of disposable income. For empirical analysis, see Keeney (1990, 1997). Lutter and Morrall (1994) provide a theoretical development and a review of the literature. Hahn, Lutter, and Viscusi (2000) provide an empirical evaluation of several regulations using health-health analysis.
of policies. Distributional issues arise, however, on both the benefit and cost sides of the ledger, and appear along a number of dimensions, including: cross-sectional (such as geographic, income, race, sector, and firm characteristics) and intertemporal (such as seasonal, annual, long term, and inter-generational). Distributional equity may be an important societal consideration, particularly in regard to impacts on people of different incomes, and two possible approaches to this issue deserve mention: distributionally weighted benefit–cost analysis and separate distributional analyses.

It is at least theoretically possible to incorporate distributional considerations into benefit–cost analysis by using a system of distributional weights, whereby greater attention is given in the analysis to the dollars received or expended by various groups in a benefit–cost analysis. This requires the specification of a set of weights, and there is neither a theoretical nor a political consensus on an appropriate set of weights. Most economists, however, do not advocate attempting to incorporate distributional considerations into benefit–cost analysis (such as via distributional weights), but recommend using separate distributional analysis as a supplement to standard benefit–cost analysis. Such distributional analysis can examine impacts on sub-groups of the population, as well as on the national distribution of income or wealth. Sub-populations that are frequently considered in policy contexts include economic sectors, government, consumers, the elderly, and children. Distributional analysis may also report on potential changes in profitability of firms, changes in employment, plant closures, changes in government revenues, and industry competitiveness.

2.2. Positive issues and analysis

Given the welfare improvements that employment of the efficiency criterion and the related assessment method of benefit–cost analysis could presumably bring to environmental policy, it is reasonable to ask what the reception has been within the three branches of the federal government—executive, legislative, and judicial—to the use of these analytical tools.

2.2.1. Executive orders

At the dawn of the modern environmental movement during the Nixon Administration in the 1970s, the Federal government “placed a high premium on immediate responses to long-neglected problems; emphasized the existence of problems rather than their

---

53 For an early treatment of the difficulty of using distributional weights to compare allocations, see Harberger (1978).
54 Some analyses have used weights based on political behavior such as tax rates. This method was suggested by Eckstein (1961). Applications include Haveman (1965) and Nwaneri (1970).
55 Examples of applications of distributional analysis to toxic waste contamination include Hird (1993), and Coates, Heid, and Munger (1994), and to air pollution include Gianessi, Peskin, and Wolff (1979) and Bingham, Anderson, and Cooley (1987).
magnitude; and often based its judgments on moral indignation directed at the behavior of those who created pollution and other risks to safety and health” (Sunstein, 2002). But, subsequently Presidents Carter, Reagan, Bush, Clinton, and Bush all introduced formal processes for reviewing economic implications of major environmental, health, and safety regulations, using Regulatory Impact Analysis. Apparently the Executive Branch, charged with designing and implementing regulations, has seen considerable need to develop a yardstick against which the efficiency of regulatory proposals can be assessed, and benefit–cost analysis has been the yardstick of choice.56

President Reagan’s Executive Order 12,291 directed executive agencies to submit any major proposed rule to the U.S. Office of Management and Budget (OMB) along with a statement assessing its regulatory impact. The order further directed that, “to the extent permitted by law,” administrative agencies were not to regulate if the costs of their regulation outweighed the benefits. Supporters of the approach emphasized that it would help achieve least-cost solutions to policy problems by bringing consistency and rationality to the administrative state, while critics contended that OMB review and benefit–cost analysis were intended not to promote efficient regulation, but simply to roll back regulation (Pildes and Sunstein, 1995).

Throughout the Reagan and Bush Administrations, Regulatory Impact Analyses (RIAs) were required under Reagan Executive Orders 12291 and 12498.57 President George H.W. Bush created a Council on Competitiveness, chaired by Vice President Quayle, which reviewed the impact on industry of selected regulations.

The Clinton Administration, like its two immediate predecessors, issued an Executive Order requiring benefit–cost analysis of all Federal regulations with expected annual costs greater than $100 million.58 Shortly after taking office in 1993, Clinton abolished the Council on Competitiveness and revoked both of the Reagan orders, replacing them with EO 12866, Regulatory Planning and Review.59 The Clinton EO was substan-

---

56 On the other hand, it should be recognized that the Office of Information and Regulatory Affairs (in the U.S. Office of Management and Budget), which reviews draft regulations and manages the process of receiving Regulatory Impact Analyses from the departments and agencies, was itself established by the Congress (through the Paperwork Reduction Act of 1980).
57 Executive Order (EO) 12291 required agencies to conduct a regulatory impact analysis for all proposed and final rules that were anticipated to have an effect on the national economy in excess of $100 million. Executive Order 12498 required, in addition, a risk assessment for all proposed and final environmental health and safety regulations. EO 12291 has been called the “foremost development in administrative law of the 1980s” (Morgenstern, 1997). The Reagan EO’s were not the first presidential effort at regulatory efficiency, however. President Nixon required a “Quality of Life” review of selected regulations in 1971, and President Ford formalized this process in EO 11281 in 1974. President Carter’s EO 12044 required analysis of proposed rules and centralized review by the Regulatory Analysis Review Group. The Administration of President George W. Bush has continued to enforce the RIA requirements of Clinton’s EO 12866, rather than issuing a new EO (Graham, 2001).
58 The threshold is not indexed for inflation and has not been modified over time. We refer to year 2000 dollars, unless we indicate otherwise.
59 In discussing Clinton’s EO 12866, many analysts also mention EO 12875, Enhancing the Intergovernmental Partnership, which limited “unfunded mandates”. While EO 12875 was part of the Administration’s regulatory reform agenda, it did not refer to the efficiency or cost-effectiveness of environmental regulations.
R.L. Revesz and R.N. Stavins

respectively and administratively similar to the Reagan orders. It was qualitatively different in
tone, however, signaling a less strict efficiency test. While the Reagan orders required
that benefits outweigh costs, the Clinton order required only that benefits justify costs.
The Clinton EO allowed that: (1) not all regulatory benefits and costs can be mone-
tized; and (2) non-monetary consequences should be influential in regulatory analysis
(Viscusi, 1993). In other ways, the Clinton EO broadened the scope of RIAs to include
“distributive impacts” and “equity” in assessing the costs and benefits of particular reg-
ulations.  

President George W. Bush kept Clinton’s executive order in place, further cementing
what was already apparent: that the use of benefit–cost analysis in the executive branch
has strong bipartisan support. This is not to say, however, that benefit–cost analysis has
become a ubiquitous part of all agency decision making. There is evidence that many
federal agencies have not complied with the executive orders to engage in meaningful
benefit–cost analysis, and the requirements for Regulatory Impact Analysis have not
necessarily improved the efficiency of individual Federal environmental rules (Hahn
and Dudley, 2004). Further, regulatory impact analysis is required only for major rules,
a small fraction of all rules issued by EPA and other agencies. Rules that do not meet
this threshold pass under the efficiency radar.

2.2.2. Legislative enactments

Over the years, Congress has sent mixed signals regarding the use of benefit–cost analy-
sis in policy evaluation. Some statutes actually require the use of benefit–cost analysis,61
whereas others have been interpreted to effectively preclude the consideration of ben-
efits and costs in the development of certain regulations.62 But this has not prevented
regulatory agencies from considering the benefits and costs of their regulatory propos-
als.63 The problem with such informal, implicit benefit–cost analysis is that it can be

---

60 “Costs and benefits shall be understood to include both quantifiable measures (to the fullest extent that
these can be usefully estimated) and qualitative measures of costs and benefits that are difficult to quantify,
but nevertheless essential to consider. Further, in choosing among alternative regulatory approaches, agencies
should select those approaches that maximize net benefits (including potential economic, environmental, pub-
lic health and safety, and other advantages; distributive impacts; and equity), unless a statute requires another
regulatory approach.” Executive Order 12866.
61 Parts of the Clean Water Act, the Consumer Product Safety Act, the Toxic Substances Control Act, the
Federal Insecticide, Fungicide, and Rodenticide Act, and the Safe Drinking Water Act explicitly allow or
require regulators to consider benefits and costs.
62 Statutes that have been interpreted (in part, at least) to restrict the ability of regulators to consider bene-
fits and costs include: the Federal Food, Drug, and Cosmetic Act; health standards under the Occupational
Safety and Health Act; safety regulations from National Highway and Transportation Safety Agency; the
Clean Air Act; the Clean Water Act; the Resource Conservation and Recovery Act; and the Comprehensive
Environmental Response, Compensation, and Liability Act.
63 There is rigorous, empirical evidence that agencies do take into account benefits and costs of regulatory
decisions, even when governing statutes do not encourage or allow such analysis to affect decisions. See, for
example: Cropper et al. (1992).
unsystematic, not subject to peer review, and carried out behind closed doors, with access limited to the particular friends of the administration. Thus, concerns arise about this approach not only on technical grounds (poor analysis), but on process grounds—it is fundamentally undemocratic.

Despite such arguments, formal benefit–cost analysis has only infrequently been used to help set the stringency of environmental standards. The body politic has favored a very different set of approaches to setting standards, such as that embraced by the Clean Air Act: set the standard to “protect the most sensitive member of the population with an adequate margin of safety.” Economists and some legal scholars have spent a great deal of time arguing that such criteria are neither reasonable nor well defined, but little change has occurred.64

In the 104th Congress, a major part of the Republicans’ “Contract with America” was a regulatory reform bill that would have made meeting a benefit–cost test a necessary condition for a broad set of regulatory actions. That bill was narrowly defeated in the Senate, and would have faced a certain Presidential veto, in any case (Sunstein, 1996).65 Subsequently, Congress considered but did not enact legislation (introduced by former Senator Fred Thompson and Senator Carl Levin) which would have required agencies to conduct (non-binding) benefit–cost analyses of new regulations and periodically of existing ones.66 While this bill never became law, the 106th Congress did pass a major piece of regulatory reform legislation, the Truth in Regulating Act (TIRA), which was signed into law by President Clinton in October 2000. The TIRA established a three-year pilot project beginning in early 2001, which required GAO to review RIAs to evaluate agencies’ benefit estimates, cost estimates, and analysis of alternative approaches, upon request by Congress. Because funding was never provided, however, TIRA was not implemented.

In addition to these attempts at cross-cutting regulatory reform, the Congresses of the Clinton years pursued efficiency within specific environmental statutes.67 In general, Congress was not successful during the 1990s at reforming individual environmen-

---

64 The significant and well known heterogeneity of costs per life saved under existing statutes (Table 1) suggests that in the absence of a benefit–cost test aimed at achieving efficiency, much could be accomplished through greater attention to simple cost-effectiveness, that is, achieving given goals or standards at minimum cost. See: Tengs et al. (1995).

65 But President Clinton did sign the Small Business Regulatory Reform Act of 1996, which provides an opportunity for the Congress to pass legislation that nullifies a regulation that does not pass a benefit–cost test (the nullification itself is then subject to possible Presidential veto, like any act of Congress).

66 Proposals for the use of a benefit–cost test for setting environmental standards have found a more receptive audience among the states. As of 1996, some 25 of 35 states surveyed reported significant environmental regulatory reform efforts, defined as including the establishment of benefit–cost criteria for promulgation of regulations (Graham and Loevzel, 1997).

67 During the 1990s, the Congress also pursued reforms of non-environmental statutes that affect environmental regulation. For example, the Accountable Pipeline Safety and Partnership Act of 1996 (104th Congress) requires the Secretary of Transportation to issue pipeline safety regulations only upon justification that benefits exceed costs.
Table 1  
Costs of selected environmental, health, and safety regulations that reduce mortality risks

<table>
<thead>
<tr>
<th>Regulation</th>
<th>Year issued</th>
<th>Agency</th>
<th>Cost per statistical life saved (millions of 2002 dollars)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logging operations</td>
<td>1994</td>
<td>OSHA</td>
<td>0.1</td>
</tr>
<tr>
<td>Unvented space hats</td>
<td>1980</td>
<td>CPSC</td>
<td>0.2</td>
</tr>
<tr>
<td>Trihalomethane drinking water standards</td>
<td>1979</td>
<td>EPA</td>
<td>0.3</td>
</tr>
<tr>
<td>Food labeling</td>
<td>1993</td>
<td>FDA</td>
<td>0.4</td>
</tr>
<tr>
<td>Passive restraints/belts</td>
<td>1984</td>
<td>NHTSA</td>
<td>0.5</td>
</tr>
<tr>
<td>Alcohol and drug control</td>
<td>1985</td>
<td>FRA</td>
<td>0.9</td>
</tr>
<tr>
<td>Seat cushion flammability</td>
<td>1984</td>
<td>FAA</td>
<td>1.0</td>
</tr>
<tr>
<td>Side-impact standards for autos</td>
<td>1990</td>
<td>NHTSA</td>
<td>1.1</td>
</tr>
<tr>
<td>Low-altitude windshear equipment and training standards</td>
<td>1988</td>
<td>FAA</td>
<td>1.8</td>
</tr>
<tr>
<td>Children’s sleepwear flammability ban</td>
<td>1973</td>
<td>CPSC</td>
<td>2.2</td>
</tr>
<tr>
<td>Benzene/fugitive emissions</td>
<td>1984</td>
<td>EPA</td>
<td>3.7</td>
</tr>
<tr>
<td>Ethylene dibromide drinking water standard</td>
<td>1991</td>
<td>EPA</td>
<td>6.0</td>
</tr>
<tr>
<td>NO\textsubscript{x} SIP Call</td>
<td>1998</td>
<td>EPA</td>
<td>6.0</td>
</tr>
<tr>
<td>Radionuclides/uranium mines</td>
<td>1984</td>
<td>EPA</td>
<td>6.9</td>
</tr>
<tr>
<td>Grain dust</td>
<td>1988</td>
<td>OSHA</td>
<td>11</td>
</tr>
<tr>
<td>Methylene chloride</td>
<td>1997</td>
<td>OSHA</td>
<td>13</td>
</tr>
<tr>
<td>Arsenic emissions standards for glass plants</td>
<td>1986</td>
<td>EPA</td>
<td>19</td>
</tr>
<tr>
<td>Arsenic emissions standards for copper smelters</td>
<td>1986</td>
<td>EPA</td>
<td>27</td>
</tr>
<tr>
<td>Hazardous waste listing for petroleum refining sludge</td>
<td>1990</td>
<td>EPA</td>
<td>29</td>
</tr>
<tr>
<td>Coke ovens</td>
<td>1976</td>
<td>OSHA</td>
<td>51</td>
</tr>
<tr>
<td>Uranium mill tailings (active sites)</td>
<td>1983</td>
<td>EPA</td>
<td>53</td>
</tr>
<tr>
<td>Asbestos/construction</td>
<td>1994</td>
<td>OSHA</td>
<td>71</td>
</tr>
<tr>
<td>Asbestos ban</td>
<td>1989</td>
<td>EPA</td>
<td>78</td>
</tr>
<tr>
<td>Hazardous waste management/wood products</td>
<td>1990</td>
<td>EPA</td>
<td>140</td>
</tr>
<tr>
<td>Sewage sludge disposal</td>
<td>1993</td>
<td>EPA</td>
<td>530</td>
</tr>
<tr>
<td>Land disposal restrictions/phase II</td>
<td>1994</td>
<td>EPA</td>
<td>2,600</td>
</tr>
<tr>
<td>Drinking water/phase II</td>
<td>1992</td>
<td>EPA</td>
<td>19,000</td>
</tr>
<tr>
<td>Formaldehyde occupational exposure limit</td>
<td>1987</td>
<td>OSHA</td>
<td>78,000</td>
</tr>
<tr>
<td>Solid waste disposal facility criteria</td>
<td>1991</td>
<td>EPA</td>
<td>100,000</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Source is Morall (2003). Only final rules are included. Estimates are from respective agencies. Non-mortality and non-health benefits were subtracted from the annual cost (numerator) to generate net cost. For each entry, the denominator is the estimated number of statistical lives saved by the regulation annually. Agency abbreviations are as follows. CPSC: Consumer Product Safety Commission; EPA: Environmental Protection Agency; NHTSA: National Highway Traffic Safety Administration; FAA: Federal Aviation Administration; FRA: Federal Railroad Administration; OSHA: Occupational Safety and Health Administration.

tal statutes, although important exceptions were the 1996 Safe Drinking Water Act (SDWA) amendments and the partial reform of pesticide permitting under the Federal Food, Drug and Cosmetic Act (FFDCA).
The 1996 SDWA Amendments include the most far-reaching requirement for benefit-cost analysis in any environmental statute. The Amendments focus EPA regulatory efforts on contaminants that pose the greatest health risks by: (1) requiring benefit-cost analysis of new rules; (2) removing the mandate that EPA regulate 25 new contaminants every three years; (3) allowing EPA to use cost information to adjust its “feasibility standards” for water system reduction of contaminants; and (4) requiring the Administrator to balance risks among contaminants to minimize the overall risk of adverse health effects (Tiemann, 1999). While the Amendments require EPA to determine whether the benefits of each new drinking water maximum contaminant level (MCL) regulation justify the costs, they also allow the Agency to adopt more stringent standards than those that maximize net benefits, explaining the reasons for not selecting the efficient standard.68

The Food Quality Protection Act of 1996 amends both the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the FFDCA, removing pesticide residues on processed food from the group of Delaney “zero-risk standard” substances. The Delaney standard has long been a target of economic criticism on the grounds that it often leads to associated costs that greatly exceed benefits.69 While the standard continues to apply to non-pesticide food additives, the Food Quality Protection Act of 1996 eliminated the distinction between pesticide residues on raw foods (which had been regulated under FFDCA section 408) and processed foods (which had been regulated under FFDCA section 409—the Delaney Clause).

It is also important to recognize several failed attempts at changes in individual statutes. Two of the environmental statutes most frequently criticized on efficiency grounds—Superfund (CERCLA) and the Clean Water Act (CWA)—remained relatively untouched by Congress in the 1990s, despite its focus on regulatory reform. Superfund’s critics have focused on the low benefits and high costs of achieving the statute’s standards (Viscusi, 1992; Breyer, 1993; Hamilton and Viscusi, 1999). Reauthorization and reform were considered during the 105th Congress, but no legislation was passed. Rather than efficiency, distributional aspects of liability issues and questions of how to finance Superfund were the major foci of legislative discussions.70 The 104th Congress pursued efficiency-oriented reform of the Clean Water Act through the reauthorization process, but the effort failed in the Senate. During the 104th Congress, the House passed a comprehensive Clean Water Act reauthorization (H.R. 961) that would have been more flexible and less prescriptive than the current statute, but the Senate did not take up the bill.

---

68 See Safe Drinking Water Act §300g-1 (4)(C). The Amendments do not allow old standards to be subjected to an ex-post benefit-cost analysis.
69 The so-called Delaney clause had the effect of forcing the Food and Drug Administration (FDA) to ban substances from human food supplies that had tested positive as animal carcinogens.
70 The taxes that support the Superfund trust fund (primarily excise taxes on petroleum and specified chemical feedstocks and a corporate environmental income tax) expired in 1995 and have not been reinstated.
Finally, it is important to note the limited effects of the legislative changes described above. The cross-cutting legislative regulatory reform measures passed in the 1990s and the efficiency-related changes to specific environmental statutes had only limited effects on regulation. This is in part due to differences between the Administration and the Congress in the acceptance of efficiency as an appropriate criterion for managing the environment and natural resources. An additional explanation is the existing statutory bias against benefit–cost analysis in some cases, particularly under the Clean Air Act. In such cases, substantial movement toward efficiency in regulation cannot be expected without substantial changes in the authorizing legislation.

The SDWA Amendments of 1996 incorporated a strong benefit–cost criterion, in comparison with other environmental statutes. However, the decisions made on MCLs since the SDWA Amendments have not placed great weight on the results of required benefit–cost analyses. Two major rules proposed since the 1996 Amendments were those regulating allowable levels of arsenic and radon in drinking water. EPA's benefit–cost analyses for the radon and arsenic MCLs can be interpreted as indicating that monetized costs exceed monetized benefits for both rules, but EPA maintained that the benefits of both rules justify their costs when unquantified benefits are included (U.S. Environmental Protection Agency, 1999).

Likewise, the regulatory reform initiatives passed by Congress in the 1990s apparently did not influence EPA's issuance of National Ambient Air Quality Standards (NAAQS) for ambient ozone and particulate matter in July, 1997. Due to their high potential compliance costs, the revised standards were immediately controversial; both the decision to tighten the standards and the quality of the research used to support the new standards came under fire. EPA's cost estimates for the ozone standard were singled out for criticism (Shogren, 1998; Lutter, 1999). On the other hand, the particulate standard exhibited expected benefits that could well exceed costs by a considerable margin.

The regulated community challenged the new NAAQS in court, and the case reached the U.S. Supreme Court in October, 2000. Under the Clean Air Act, EPA is required to set health-based standards for specified pollutants without consideration of abatement costs. The Supreme Court ruled unanimously in February, 2001, that the Clean Air Act does not allow EPA to consider costs in setting NAAQS for the air pollutants in question (and that the statute’s mandate that the NAAQS protect the public health with “an adequate margin of safety” allows an acceptable scope of discretion to EPA).

Overall, the differences in opinion between Congress and the executive branch (especially EPA) on the usefulness of efficiency analysis have resulted in an effective stalemate. Even where statutes have been explicitly altered to require benefit–cost analysis, as was the case for the setting of MCLs under the Safe Drinking Water Act, rules

---

71 The arsenic rule was finalized on January 22, 2001, but implementation was delayed while the rule was taken under review by the George W. Bush Administration, citing concerns about the rule’s costs and benefits. After an expedited review by the National Academy of Sciences, in October, 2001, EPA Administrator Whitman announced the Agency’s intention to enforce the Clinton arsenic standard.
promulgated during the 1990s were not any more or less efficient than rules promulgated during earlier decades. On the other hand, Congressional efforts at generic “regulatory reform” are unlikely to disappear from the policy landscape, and there will continue to be attempts—sometimes successful—to introduce benefit–cost tests into individual environmental statutes.

2.2.3. Judicial recognition

The Federal judiciary also plays a key role in furthering the use of analytical methodologies to assist agency decision making. As Sunstein (2002) notes, over the years courts have implemented a series of benefit–cost “default rules” to deal with Congressional silences and ambiguities. These default rules, while not part of administrative law doctrine, impute to Congressional silence an intent to permit (and perhaps even require) administrative agencies to consider regulatory costs when issuing regulations. The default rules reflect a widespread judicial acceptance of benefit–cost analysis in regulatory rulemaking.72

Notably, Justice Stephen Breyer of the U.S. Supreme Court has advocated for more aggressive background rules with respect to risk-risk analysis. In the seminal case of American Trucking v. Whitman,73 referenced above, Justice Breyer argued for a general presumption requiring agencies to engage in benefit–cost analyses when regulating, noting that these analyses would also necessarily involve assessments of risk tradeoffs. In this, he joined ranks with economists and legal scholars who have argued for a judicial presumption in favor of benefit–cost and risk-tradeoff analyses. Breyer’s concurrence marks the arrival of risk tradeoff analysis—and health-health tradeoff analysis, in particular—in the Supreme Court and paves the way for future challenges based on such tradeoffs (Rascoff and Revesz, 2002).

2.2.4. A political economy perspective on how standards are set

Granting the merits and relatively widespread acceptance of analytical methods for assessing the tradeoffs inherent in environmental regulation, why has the use of analytical techniques not become more common in environmental policy? Why instead has Congress continued to legislate frequently without regard to benefits and costs? This section reviews positive political economy accounts of how environmental standards are set.

First, some regulations permit established firms to extract rents and establish barriers to entry that convey to them a competitive advantage (Keohane, Revesz, and Stavins, 1998). This is consistent with the empirical reality that the impetus for regulation often comes, either explicitly or implicitly, from regulated firms themselves. For example, a

command-and-control standard that limits a firm’s aggregate emissions may cause firms to reduce their output to meet the environmental requirement. This output restriction can push the price of a firm’s product above its average cost, and as a result, the firm can earn rents. Vigorous competition would dissipate this rent, but environmental regulations often create barriers to entry by imposing stringent pollution standards on new sources, thus giving a significant competitive advantage to established polluters (Maloney and McCormick, 1982).

Second, some industries enjoy strong economies of scale and prefer uniform federal regulation to a patchwork of potentially more efficient state standards. Indeed, having to manufacture different products for sale in different states can destroy the advantages related to economies of scale. National uniformity can come at the expense of regulations more narrowly tailored to achieve optimal outcomes.

Third, even if environmental regulation does not raise the profits of an entire industry, it can benefit certain firms within an industry (Keohane, Revesz, and Stavins, 1998). Firms within an industry likely will incur different costs in the regulatory requirements, because some firms will be able to adjust their production processes more easily than others. These relative beneficiaries of government regulation are thus likely to oppose relaxing regulatory requirements, and may even favor extending them.

Fourth, the impetus for regulation sometimes comes from manufacturers of pollution control equipment, environmentally friendly technologies, or inputs to production processes favored by the regulatory regime. For example, firms specializing in the cleanup of hazardous waste sites emerged in response to the Federal Superfund statute. Similarly, the ethanol industry has strongly supported stricter regulation of gasoline. As a result of its efforts, the Clean Air Act’s clean fuels program provides strong incentives for the use of ethanol, and the Federal government has provided large subsidies to ethanol producers.

Fifth, environmental regulation often imposes disproportionate costs on some regions of the country. Regions that incur lower than average costs from regulation become comparatively more attractive to mobile capital, which may bring economic benefits such as jobs and tax revenues. Such regions sometimes push for Federal regulation that will impose disproportionate costs on other regions (Pashigian, 1985).

3. Choosing instruments: the means of environmental policy

Environmental policies typically combine the identification of a goal with some means to achieve that goal. In section 2 of this chapter, we examined the criteria and methods that economics can bring to bear on the choice of targets. In this section, we examine the means—the instruments—that can be employed by governments to achieve given policy objectives. We begin with normative issues and then turn to positive analysis.
3.1. Normative issues and analysis

Even if the goals and targets of environmental policies are taken as given, economic analysis can bring valuable insights to the assessment and design of environmental policies. We begin by considering criteria that can be brought to bear on the search for better policy instruments, and then turn to an enumeration of major categories of environmental policy instruments, including both conventional, command-and-control and the newer breed of market-based instruments. Cross-cutting issues are considered, including uncertainty, technological change, and distributional issues. We examine lessons that emerge from research and experience.

3.1.1. Potential criteria for choosing among policy instruments

A variety of criteria have been posited as relevant for choosing environmental policy instruments, including: (1) will the policy instrument achieve the stated goal or standard; (2) will it do so at the lowest possible cost, including both private-sector compliance and public-sector monitoring and enforcement; (3) will it provide government with the information it needs to implement the policy; (4) will the instrument be flexible in the face of changes in tastes and technology; (5) will the instrument provide dynamic incentives for research, development, and adoption of better pollution-abatement technologies; (6) will the implementation of the policy instrument result in an equitable distribution of the benefits and costs of environmental protection; and (7) will the policy be politically feasible in terms of enactment and implementation? Items (1) through (5) together refer to a comprehensive notion of the criterion of cost-effectiveness, while item (6) refers to distributional equity, and item (7) refers to political feasibility.\(^74\)

First, to be more precise, by cost-effectiveness we mean that allocation of control among sources that results in the aggregate target being achieved at the lowest possible cost, that is, the allocation which satisfies the following cost-minimization problem:

\[
\min_{\{r_i\}} C = \sum_{i=1}^{N} c_i(r_i) \\
\text{s.t. } \sum_{i=1}^{N} [u_i - r_i] \leq E \\
\text{and } 0 \leq r_i \leq u_i
\]

\(^74\) This list originated with Bohm and Russell (1985). As indicated above, we include the first potential criterion—environmental effectiveness—in a comprehensive definition of cost-effectiveness, but it can also be considered on its own. For example, it has been argued that in some cases the use of market-based instruments has made it politically and/or economically feasible to achieve more stringent goals than otherwise possible (Ellerman et al., 2000; Ellerman, 2003; Harrison, 2003).
where \( r_i \) = reductions in emissions (abatement or control) by source \( i \) (\( i = 1 \) to \( N \));
\( c_i(r_i) \) = cost function for source \( i \);
\( C \) = aggregate cost of control;
\( u_i \) = uncontrolled emissions by source \( i \); and
\( E \) = the aggregate emissions target imposed by the regulatory authority.

If the cost functions are convex, then necessary and sufficient conditions for satisfaction of the constrained optimization problem posed by equations (10) through (12) are the following, among others (Kuhn and Tucker, 1951):

\[
\frac{\partial c_i(r_i)}{\partial r_i} - \lambda \geq 0
\]

(13)

\[
r_i \cdot \left[ \frac{\partial c_i(r_i)}{\partial r_i} - \lambda \right] = 0
\]

(14)

Equations (13) and (14) together imply the crucial condition for cost-effectiveness that all sources (that exercise some degree of control) experience the same marginal abatement costs (Baumol and Oates, 1988). Thus, when examining alternatives types of environmental policy instruments, a key question is whether particular instruments are likely to result in marginal abatement costs being equated across sources.\(^75\)

3.1.2. Alternative policy instruments

The most frequently employed delineation of environmental policy instruments is that of command-and-control versus market-based approaches. Conventional approaches to regulating the environment—frequently characterized as command-and-control\(^76\)—allow relatively little flexibility in the means of achieving goals. Such policy instruments tend to force firms to take on similar shares of the pollution-control burden, regardless of the cost, sometimes by setting uniform standards for firms, the most prevalent of which are technology- and performance-based standards.\(^77\)

Market-based instruments encourage behavior through market signals, rather than through explicit directives regarding pollution control levels or methods. These policy

---

75 For purposes of clarity, the model of cost-effectiveness, above, and subsequent models of specific policy instruments refer to uniformly-mixed flow pollutants. Little additional insight is gained but much is sacrificed in terms of transparency and tractability by modeling more complex non-uniformly mixed stock pollutants. Where the results are not robust to this simplification, we recognize the complexities in the text.

76 The phrase “command-and-control” is by far the most commonly employed characterization for conventional environmental policy instruments, including uniform performance and technology standards. Admittedly, the phrase has an inescapable negative stigma associated with it, and so a better, more neutral description of this category of policy approaches might be “prescriptive instruments.” But because “command-and-control” is the generally accepted name for this category, we employ it in this chapter.

77 Note that uniform standards can specify the amount of pollution that can be released into the environment (emission standard) or the permissible concentration of pollution in the air, water, or soil (ambient standard). The cost-effective allocation consistent with ambient standards requires equalization of the marginal costs to reduce a unit of ambient concentration, rather than emission.
instruments can reasonably be described as “harnessing market forces,” because if they are well designed and properly implemented, they encourage firms or individuals to undertake pollution control efforts that are in their own interests and that collectively meet policy goals. Market-based instruments fall within four categories: pollution charges, tradeable permits, market-friction reductions, and government subsidy reductions. Liability rules can also be thought of as a market-based instrument, because they provide incentives for firms to take into account the potential environmental damages of their decisions, allowing full flexibility in technology and control practices (Revesz, 1997c).

3.1.2.1. Command-and-control versus market-based instruments

Market-based instruments offer the potential for dynamic cost-effectiveness, but problems may arise in translating theory into practice (Hahn and Axtell, 1995), and it has been difficult to measure the magnitude of the gains of moving from command-and-control to incentive-based mechanisms. One frequently-cited survey of eleven empirical studies of air pollution control found that the ratio of actual, aggregate costs of the conventional (command-and-control) approach to the aggregate costs of least-cost benchmarks ranged from 1.07 for sulfate emissions in the Los Angeles area to 22.0 for hydrocarbon emissions at all domestic DuPont plants (Tietenberg, 1985). It is important not to misinterpret these numbers, however, since actual, command-and-control instruments were essentially contrasted with theoretical benchmarks of cost-effectiveness, that is, what a perfectly functioning market-based instrument would achieve in theory. A more useful comparison among policy instruments might involve either idealized versions of both market-based systems and alternatives, or—better yet—realistic versions of both (Hahn and Stavins, 1992).

Where there is significant heterogeneity of costs, command-and-control methods will not be cost-effective. Holding all firms to the same target will be unduly expensive, because it fails to recognize abatement cost heterogeneity. In reality, costs can vary enormously due to production design, physical configuration, age of assets, and other factors. For example, the marginal costs of controlling lead emissions have been estimated to range from $13 to $56,000 per ton (Hartman, Wheeler, and Singh, 1994; Morgenstern, 2000). But where costs are similar among sources, command-and-control

---


79 Other taxonomies of regulatory instruments are possible, and some take a more inclusive view, including—for example—contractual approaches. On this, see Menell (2002).

80 In other cases, researchers have contrasted hypothetical costs of a CAC program with the actual compliance costs associated with the use of a market-based instrument (Keohane, 2003).

81 Harrington and Morgenstern (2003) attempt to do this by comparing actual experiences in Europe and the United States with market-based and conventional policy instruments.
instruments may perform equivalent to (or better than) market-based instruments, depending on transactions costs, administrative costs, possibilities for strategic behavior, political costs, and the nature of the pollutants (Newell and Stavins, 2003).82

In theory, if properly designed and implemented, market-based instruments allow any desired level of pollution cleanup to be realized at the lowest overall cost to society, by providing incentives for the greatest reductions in pollution by those firms that can achieve the reductions most cheaply. Rather than equalizing pollution levels among firms, market-based instruments equalize their marginal abatement costs (Montgomery, 1972; Baumol and Oates, 1988; Tietenberg, 1995). Command-and-control approaches could—in theory—achieve this cost-effective solution, but this would require that different standards be set for each pollution source, and, consequently, that policy makers obtain detailed information about the compliance costs each firm faces. Such information is simply not available to government. By contrast, market-based instruments provide for a cost-effective allocation of the pollution control burden among sources without requiring the government to have this information.

In addition, market-based instruments have the potential to bring down abatement costs over time (that is, to be dynamically cost effective) by providing incentives for companies to adopt cheaper and better pollution-control technologies. This is because with market-based instruments, most clearly with emission taxes, it pays firms to clean up a bit more if a sufficiently low-cost method (technology or process) of doing so can be identified and adopted (Downing and White, 1986; Ellerman, 2003; Maloney, 1989; Milliman and Prince, 1989; Jaffe and Stavins, 1995; Carlson et al., 2000; Popp, 2002; Keohane, 2001; Tietenberg, 2003). However, the ranking among policy instruments, in terms of their respective impacts on technology innovation and diffusion, is not unequivocal (Jaffe, Newell, and Stavins, 2003).

3.1.2.2. Pollution charges Pollution charge systems assess a fee or tax on the amount of pollution that firms or sources generate (Pigou, 1920). Consequently, it is worthwhile for firms to reduce emissions to the point where their marginal abatement costs are equal to the common tax rate.83 By definition, actual emissions are equal to unconstrained emissions minus emissions reductions, that is, \( e_i = u_i - r_i \). A source’s cost minimization problem in the presence of an emissions tax, \( t \), is given by:

\[
\begin{align*}
\min_{\{r_i\}} & \quad c_i(r_i) + t \cdot (u_i - r_i) \\
\text{s.t.} & \quad r_i \geq 0
\end{align*}
\]

The result for each source is:

\[
\frac{\partial c_i(r_i)}{\partial r_i} - t \geq 0
\]

82 Also see: Atkinson and Lewis (1974); Spofford (1984); and Maloney and Yandle (1984).

83 For an examination of the robustness of this result in the presence of non-competitive conditions, see Cropper and Oates (1992).
Equations (17) and (18) imply that each source (that exercises a positive level of control) will carry out abatement up to the point where its marginal control costs are equal to the tax rate. Hence, marginal abatement costs will be equated across sources, satisfying the condition for cost-effectiveness specified by equations (13) and (14).

A challenge with charge systems is identifying the appropriate tax rate. For social efficiency, it should be set equal to the marginal benefits of cleanup at the efficient level of cleanup, but policy makers are more likely to think in terms of a desired level of cleanup, and they do not know beforehand how firms will respond to a given level of taxation. An additional problem posed by pollution taxes is associated with their distributional consequences for regulated sources. Despite the fact that such systems minimize aggregate social costs, these systems may be *more* costly than comparable command-and-control instruments for regulated firms. This is because with the tax approach, firms pay both their abatement costs plus taxes on their residual emissions. For the calculation of aggregate costs in a social benefit–cost or cost-effectiveness analysis, tax payments are simply transfers, and so are excluded from the calculations.

The conventional wisdom is that charge systems have been ignored in the United States, but this is not really correct. If one defines charge systems broadly, a significant number of applications can be identified (Stavins, 2003). The closest that any U.S. charge systems come to operating as true Pigovian taxes may be the increasingly common *unit-charge* systems for financing municipal solid waste collection, where households and businesses are charged the incremental costs of collection and disposal. So-called “pay-as-you-throw” policies, where users pay in proportion to the volume of their waste, are now used in well over one thousand jurisdictions. The collective experience provides evidence that unit charges have been successful in reducing the volume of household waste generated.84

Another important set of charge systems implemented in the United States has been *deposit refund systems*, whereby consumers pay a surcharge when purchasing potentially polluting products, and receive a refund when returning the product to an approved center for recycling or proper disposal. A number of states have implemented this approach through “bottle bills” to control litter from beverage containers and to reduce the flow of solid waste to landfills (Bohm, 1981; Menell, 1990), and the concept has also been applied to lead-acid batteries (Table 2).

In addition, there has been considerable use of *environmental user charges* in the United States, through which specific environmentally related services are funded (Table 3). Examples include *insurance premium taxes* (Table 4), such as those formerly used to fund partially the clean-up of hazardous waste sites through the Superfund

---

84 See: McFarland (1972); Wertz (1976); Stevens (1978); Efaw and Lanen (1979); Skumatz (1990); Lave and Gruenspecht (1991); Repetto et al. (1992); Miranda et al. (1994); Fullerton and Kinnaman (1996); and Menell (2003).
Table 2
Deposit-refund systems for two regulated products

<table>
<thead>
<tr>
<th>State</th>
<th>Year of initiation</th>
<th>Amount of deposit ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specified beverage containers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oregon</td>
<td>1972</td>
<td>0.05a</td>
</tr>
<tr>
<td>Vermont</td>
<td>1973</td>
<td>0.05</td>
</tr>
<tr>
<td>Maine</td>
<td>1978</td>
<td>0.05</td>
</tr>
<tr>
<td>Michigan</td>
<td>1978</td>
<td>0.10</td>
</tr>
<tr>
<td>Iowa</td>
<td>1979</td>
<td>0.05</td>
</tr>
<tr>
<td>Connecticut</td>
<td>1980</td>
<td>0.05</td>
</tr>
<tr>
<td>Delaware</td>
<td>1983</td>
<td>0.05</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>1983</td>
<td>0.05</td>
</tr>
<tr>
<td>New York</td>
<td>1983</td>
<td>0.05</td>
</tr>
<tr>
<td>California</td>
<td>1987</td>
<td>0.025–0.06b</td>
</tr>
<tr>
<td>Auto batteries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minnesota</td>
<td>1988</td>
<td>5.00</td>
</tr>
<tr>
<td>Maine</td>
<td>1989</td>
<td>10.00</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>1989</td>
<td>5.00</td>
</tr>
<tr>
<td>Washington</td>
<td>1989</td>
<td>5.00</td>
</tr>
<tr>
<td>Arizona</td>
<td>1990</td>
<td>5.00</td>
</tr>
<tr>
<td>Connecticut</td>
<td>1990</td>
<td>5.00</td>
</tr>
<tr>
<td>Michigan</td>
<td>1990</td>
<td>6.00</td>
</tr>
<tr>
<td>Idaho</td>
<td>1991</td>
<td>5.00</td>
</tr>
<tr>
<td>New York</td>
<td>1991</td>
<td>5.00</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>1991</td>
<td>5.00</td>
</tr>
<tr>
<td>Arkansas</td>
<td>1991</td>
<td>10.00</td>
</tr>
</tbody>
</table>

a$0.02 for refillable containers.
bDeposits depend upon materials and size of containers.

program (Barthold, 1994). Another set of environmental charges are sales taxes on motor fuels, ozone-depleting chemicals, agricultural inputs, and low-mileage motor vehicles (Table 5). Finally, tax differentiation has become part of a considerable number of Federal and state attempts to encourage the use of renewable energy sources (Table 6).

3.1.2.3. Tradeable permit systems Tradeable permits—in theory—can achieve the same cost-minimizing allocation of the control burden as a charge system, while

85 The taxes that previously supported the Superfund trust fund—primarily excise taxes on petroleum and specified chemical feedstocks and a corporate environmental income tax—expired in 1995, and have not been reinstated.

86 Thirty years ago, Crocker (1966) and Dales (1968) independently developed the idea of using transferable discharge permits to allocate the pollution-control burden among sources. Montgomery (1972) provided the
### Table 3
Federal user charges

<table>
<thead>
<tr>
<th>Item taxed</th>
<th>First enacted/ modified</th>
<th>Rate</th>
<th>Use of revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trucks and trailers (excise tax)</td>
<td>1917/1984</td>
<td>12%</td>
<td>Highway Trust Fund/Mass Transit Account</td>
</tr>
<tr>
<td>Sport fishing equipment</td>
<td>1917/1984</td>
<td>10% (except 3% for out board motors)</td>
<td>Sport Fishing Restoration Account of Aquatic Resources Trust Fund</td>
</tr>
<tr>
<td>Firearms and ammunition</td>
<td>1918/1969</td>
<td>10%</td>
<td>Federal Aid to Wildlife Program</td>
</tr>
<tr>
<td>Noncommercial motorboat fuels</td>
<td>1932–1992</td>
<td>$1.83/gal</td>
<td>Aquatic Resource Trust Fund</td>
</tr>
<tr>
<td>Motor fuels</td>
<td>1932/1993</td>
<td>$1.83/gal</td>
<td>Highway Trust Fund/Mass Transit Account</td>
</tr>
<tr>
<td>Non-highway recreational fuels &amp;</td>
<td>1932/1993</td>
<td>$1.83/gal gasoline $2.43/gal diesel</td>
<td>National Recreational Trails Trust Fund and Wetlands Account of Aquatic Resources Trust Fund</td>
</tr>
<tr>
<td>small-engine motor fuels</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Annual use of heavy vehicles</td>
<td>1951/1993</td>
<td>$100–$500/vehicle</td>
<td>Highway Trust Fund/Mass Transit Account</td>
</tr>
<tr>
<td>Bows and arrows</td>
<td>1972/1984</td>
<td>11%</td>
<td>Federal Aid to Wildlife Program</td>
</tr>
<tr>
<td>Inland waterways fuels</td>
<td>1978/1993</td>
<td>$2.33/gal</td>
<td>Inland Waterways Trust Fund</td>
</tr>
</tbody>
</table>


### Table 4
Federal insurance premium taxes

<table>
<thead>
<tr>
<th>Item or action taxed</th>
<th>First enacted/ modified</th>
<th>Rate</th>
<th>Use of revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal production</td>
<td>1977/1987</td>
<td>$1.10/ton underground; $0.55/ton surface</td>
<td>Black Lung Disability Trust Fund</td>
</tr>
<tr>
<td>Chemical production</td>
<td>1980/1986</td>
<td>$0.22 to $4.88/ton</td>
<td>Superfund (CERCLA)</td>
</tr>
<tr>
<td>Petroleum production</td>
<td>1980/1986</td>
<td>$0.097/barrel crude</td>
<td></td>
</tr>
<tr>
<td>Corporate income</td>
<td>1986</td>
<td>0.12% of alternative minimum taxable income over $2 million</td>
<td></td>
</tr>
<tr>
<td>Petroleum-based fuels, except propane</td>
<td>1986/1990 (expired 1995)</td>
<td>$0.001/gal</td>
<td>Leaking Underground Storage Trust Fund</td>
</tr>
<tr>
<td>Petroleum and petroleum products</td>
<td>1989/1990</td>
<td>$0.05/barrel</td>
<td>Oil Spill Liability Trust Fund</td>
</tr>
</tbody>
</table>

Table 5
Federal sales taxes

<table>
<thead>
<tr>
<th>Item or action taxed</th>
<th>First enacted/modified</th>
<th>Rate</th>
<th>Use of revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>New tires</td>
<td>1918/1984</td>
<td>$0.15–$0.50/pound</td>
<td>U.S. Treasury</td>
</tr>
<tr>
<td>New automobiles exceeding fuel efficiency standards</td>
<td>1978/1990</td>
<td>$1,000–$7,700 per auto</td>
<td>U.S. Treasury</td>
</tr>
</tbody>
</table>


Table 6
Federal tax differentiation

<table>
<thead>
<tr>
<th>Item or action taxed</th>
<th>Provision</th>
<th>First enacted/modified</th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor fuels excise tax exemptions</td>
<td>Natural gas</td>
<td>1978/1990</td>
<td>$.07/gal</td>
</tr>
<tr>
<td></td>
<td>Methanol</td>
<td>1978/1990</td>
<td>$.06/gal</td>
</tr>
<tr>
<td></td>
<td>Ethanol</td>
<td>1978/1990</td>
<td>$.054/gal</td>
</tr>
<tr>
<td>Income tax credits</td>
<td>Alcohol fuels</td>
<td>1980/1990</td>
<td>$.60/gal; $0.54/gal ethanol</td>
</tr>
<tr>
<td></td>
<td>Business energy</td>
<td>1980/1990</td>
<td>10% solar; 10% geothermal</td>
</tr>
<tr>
<td></td>
<td>Non-conventional fuels</td>
<td>1980/1990</td>
<td>$3.00/Btu-barrel equivalent of oil</td>
</tr>
<tr>
<td></td>
<td>Wind production</td>
<td>1992</td>
<td>1.5¢/kWh</td>
</tr>
<tr>
<td></td>
<td>Biomass production</td>
<td>1992</td>
<td>1.5¢/kWh</td>
</tr>
<tr>
<td></td>
<td>Electric automobiles</td>
<td>1992</td>
<td>10% credit</td>
</tr>
<tr>
<td>Other income tax provisions</td>
<td>Van pools</td>
<td>1978</td>
<td>Tax-free employer provided benefits</td>
</tr>
<tr>
<td></td>
<td>Mass transit passes</td>
<td>1984/1992</td>
<td>Tax-free employer provided benefits</td>
</tr>
<tr>
<td></td>
<td>Utility rebates</td>
<td>1992</td>
<td>Exclusion of subsidies from utilities for energy conservation measures</td>
</tr>
<tr>
<td>Tax exempt private activity bonds</td>
<td>Mass transit</td>
<td>1968/1986</td>
<td>Interest exempt from Federal taxation</td>
</tr>
<tr>
<td></td>
<td>Sewage treatment</td>
<td>1968/1986</td>
<td>Interest exempt from Federal taxation</td>
</tr>
<tr>
<td></td>
<td>Solid waste disposal</td>
<td>1968/1986</td>
<td>Interest exempt from Federal taxation</td>
</tr>
<tr>
<td></td>
<td>Water treatment</td>
<td>1968/1986</td>
<td>Interest exempt from Federal taxation</td>
</tr>
<tr>
<td></td>
<td>High speed rail</td>
<td>1988/1993</td>
<td>Interest exempt from Federal taxation</td>
</tr>
</tbody>
</table>


\(^a\)Exemptions from the motor fuels excise tax of $0.183/gallon (see Table 3).

first rigorous proof that such a system could provide a cost-effective policy instrument. A sizeable literature has followed, much of it stemming from Hahn and Noll (1982). Early surveys were provided by Tietenberg
avoiding the problems of uncertain responses by firms and the distributional consequences of taxes. Under a tradable permit system, an allowable overall level of pollution, $E$, is established, and allocated among firms in the form of permits. Firms that keep their emission levels below their allotted level may sell their surplus permits to other firms or use them to offset excess emissions in other parts of their operations.

Let $q_{0i}$ be the initial allocation of emission permits to source $i$, such that:

$$\sum_{i=1}^{N} q_{0i} = E$$ (19)

Then, if $p$ is the market-determined price of tradeable permits, a single firm’s cost minimization problem is given by:

$$\min_{\{r_i\}} \left[ c_i(r_i) + p \cdot (u_i - r_i - q_{0i}) \right]$$ (20)

s.t. $r_i \geq 0$ (21)

The result for each source is:

$$\frac{\partial c_i(r_i)}{\partial r_i} - p \geq 0$$ (22)

$$r_i \cdot \left[ \frac{\partial c_i(r_i)}{\partial r_i} - p \right] = 0$$ (23)

Equations (22) and (23) together imply that each source (that exercises a positive level of control) will carry out abatement up to the point where its marginal control costs are equal to the market-determined permit price. Hence, the environmental constraint, $E$, is satisfied, and marginal abatement costs are equated across sources, satisfying the condition for cost-effectiveness. Note that the unique cost-effective equilibrium is achieved independent of the initial allocation of permits (Montgomery, 1972). This is of great importance politically, as we discuss below in section 3.2.

(1980, 1985). Much of the literature may be traced to Coase’s (1960) treatment of negotiated solutions to externality problems. As indicated previously, the simple model posited above, as well as the prior model of emission taxes, assumes the existence of a uniformly-mixed pollutant, in which case the focus of regulation can be exclusively on emissions, as opposed to ambient concentrations. There is a sizable literature that explores tradeable permit and other policy instruments in the context of non-uniformly-mixed pollution problems. See, for example: Montgomery (1972); and Nash and Revesz (2001).

87 This assumes that the allocation is made without charge, but it could also be through sale or auction, in which case the distributional implications of a comparable tradeable permit program are similar to the emission tax previously described. Likewise, a revenue-neutral emissions tax, in which revenues are refunded to regulated firms (but not in proportion to their emissions levels), can resemble—in distributional terms—a comparable tradeable permit program in which the permits are allocated without charge.

88 The simple program described above is a “cap-and-trade” system, but some systems operate as “credit programs,” where permits or credits are assigned only when a source reduces emissions below what is required by source-specific limits.

89 This is true unless particularly perverse types of transactions costs are present (Stavins, 1995).
Table 7

Major U.S. tradeable permit systems

<table>
<thead>
<tr>
<th>Program</th>
<th>Traded commodity</th>
<th>Period of operation</th>
<th>Environmental and economic effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emissions Trading Program</td>
<td>Criteria air pollutants under the Clean Air Act</td>
<td>1974–present</td>
<td>Environmental performance unaffected; total savings of $5–12 billion</td>
</tr>
<tr>
<td>Leaded Gasoline Phasedown</td>
<td>Rights for lead in gasoline among refineries</td>
<td>1982–1987</td>
<td>More rapid phaseout of leaded gasoline; $250 million annual savings</td>
</tr>
<tr>
<td>Water Quality Trading</td>
<td>Point-nonpoint sources of nitrogen &amp; phosphorus</td>
<td>1984–1986</td>
<td>No trading occurred, because ambient standards not binding</td>
</tr>
<tr>
<td>CFC Trading for Ozone Protection</td>
<td>Production rights for some CFCs, based on depletion potential</td>
<td>1987–present</td>
<td>Environmental targets achieved ahead of schedule; effect of TP system unclear</td>
</tr>
<tr>
<td>Heavy Duty Engine Trading</td>
<td>Averaging, banking, and trading of credits for NOx and particulate emissions</td>
<td>1992–present</td>
<td>Standards achieved; cost savings unknown</td>
</tr>
<tr>
<td>Acid Rain Reduction</td>
<td>SO2 emission allowances; mainly among electric utilities</td>
<td>1995–present</td>
<td>SO2 reductions achieved ahead of schedule; annual savings of $1 billion per year</td>
</tr>
<tr>
<td>RECLAIM Program</td>
<td>SO2 and NOx emissions by large stationary sources</td>
<td>1994–present</td>
<td>Unknown</td>
</tr>
<tr>
<td>Northeast Ozone Transport</td>
<td>Primarily NOx emissions by large stationary sources</td>
<td>1999–present</td>
<td>Unknown</td>
</tr>
</tbody>
</table>


In theory, a number of factors can adversely affect the performance of a tradeable permit system, including: concentration in the permit market (Hahn, 1984; Misolek and Elder, 1989); concentration in the product market (Maleug, 1990); transaction costs (Stavins, 1995); non-profit maximizing behavior, such as sales or staff maximization (Tschirhart, 1984); the preexisting regulatory environment (Bohi and Burtraw, 1992); and the degree of monitoring and enforcement (Keeler, 1991; and Montero, 2003).

Tradeable permits have been the most frequently used market-based system in the United States (U.S. Environmental Protection Agency, 2000a; Tietenberg, 1997a). A selection of programs is summarized in Table 7. The U.S. EPA first experimented with emissions trading in 1974, as part of the Clean Air Act’s program for improving local air quality, and later codified these initiatives in its Emissions Trading Program in 1986 (Tietenberg, 1985; Hahn, 1989; Foster and Hahn, 1995). Significant applications include: EPA’s emissions trading program (Tietenberg, 1985; Hahn, 1989); the leaded gasoline phasedown; water quality permit trading (Hahn, 1989; Stephenson, Norris, and Shabman, 1998); CFC trading (Hahn and McGartland, 1989); the sulfur dioxide (SO2) allowance trading system for acid rain control; the RECLAIM program in the Los Angeles metropolitan region (Harrison, 1999); and tradeable devel-
development rights for land use. At least two of these programs—lead trading and the SO$_2$ allowance system—merit further comment.

The purpose of the lead trading program, developed in the 1980s, was to allow gasoline refiners greater flexibility in meeting emission standards at a time when the lead-content of gasoline was reduced to 10 percent of its previous level. In 1982, EPA authorized inter-refinery trading of lead credits, a major purpose of which was to lessen the financial burden on smaller refineries, which were believed to have significantly higher compliance costs. If refiners produced gasoline with a lower lead content than was required, they earned lead credits. In 1985, EPA initiated a program allowing refineries to bank lead credits, and subsequently firms made extensive use of this option. In each year of the program, more than 60 percent of the lead added to gasoline was associated with traded lead credits (Hahn and Hester, 1989), until the program was terminated at the end of 1987, when the lead phasedown was completed.

The lead program was successful in meeting its environmental targets, although it may have produced some temporary geographic shifts in use patterns (Anderson, Hofmann, and Rusin, 1990). Although the benefits of the trading scheme are more difficult to assess, the level of trading activity and the rate at which refiners reduced their production of leaded gasoline suggest that the program was relatively cost-effective (Kerr and Maré, 1997; Nichols, 1997). The high level of trading among firms far surpassed levels observed in earlier environmental markets. EPA estimated savings from the lead trading program of approximately 20 percent below alternative programs that did not provide for lead banking, a cost savings of about $250 million per year (U.S. Environmental Protection Agency, Office of Policy Analysis, 1985). Furthermore, the program appears to have provided measurable incentives for cost-saving technology diffusion (Kerr and Newell, 2003).

The most important application made to date of a market-based instrument for environmental protection has been the SO$_2$ allowance trading program for acid rain control, established under the Clean Air Act Amendments of 1990, and intended to reduce SO$_2$ emissions by 10 million tons below 1980 levels (Ferrall, 1991). A robust market of bilateral SO$_2$ permit trading gradually emerged, resulting in cost savings on the order of $1 billion annually, compared with the costs under some command-and-control regulatory alternatives (Carlson et al., 2000). Although the program had

---

90 In addition, the Energy Policy and Conservation Act of 1975 established Corporate Average Fuel Economy (CAFE) standards for automobiles and light trucks, requiring manufacturers to meet minimum sales-weighted average fuel efficiency for their fleets sold in the United States. A penalty is charged per car sold per unit of average fuel efficiency below the standard. The program operates like an intra-firm tradeable permit system, since manufacturers can undertake efficiency improvements wherever they are cheapest within their fleets. For reviews of the program’s costs relative to “equivalent” gasoline taxes, see: Crandall et al. (1986); Goldberg (1998); and National Research Council (2002). Light trucks, which are defined by the Federal government to include “sport utility vehicles,” face weaker CAFE standards.

91 The choice of counterfactual for purposes of comparison in such estimates of cost savings is important. The estimate above represents a cost savings of about 30 percent. Employing a different counterfactual for comparison, Keohane (2003) estimates cost savings between 15 and 25 percent.
low levels of trading in its early years (Burtraw, 1996), trading increased significantly over time (Schmalensee et al., 1998; Stavins, 1998; Burtraw and Mansur, 1999; Ellerman et al., 2000).

Concerns were expressed early on that state regulatory authorities would hamper trading in order to protect their domestic coal industries, and some research indicates that state public utility commission cost-recovery rules provided poor guidance for compliance activities (Rose, 1997; Bohi, 1994). Other analysis suggests that this was not a major problem (Bailey, 1996). Similarly, in contrast to early assertions that the structure of EPA's small permit auction market would cause problems (Cason, 1995), the evidence now indicates that this had little or no effect on the vastly more important bilateral trading market (Joskow, Schmalensee, and Bailey, 1998).

The reduction of emissions through the allowance trading program apparently has had exceptionally positive welfare effects, with benefits being as much as six times greater than costs (Burtraw et al., 1998). The large benefits of the program are due mainly to the positive human health impacts of decreased local SO2 and particulate concentrations, not the ecological impacts of reduced long-distance transport of acid deposition. This contrasts with what was understood and assumed at the time of the program's enactment in 1990.

3.1.2.4. Market friction reduction Market friction reduction can also serve as a policy instrument for environmental protection. Market creation establishes markets for inputs or outputs associated with environmental quality. Finally, since well-functioning markets depend on the existence of well-informed producers and consumers, information programs can help foster market-oriented solutions to environmental problems. Product labeling requirements can improve the information set available to consumers, as can various types of reporting requirements.

One prominent example of market creation is provided by measures that facilitate the voluntary exchange of water rights and thus promote more efficient allocation and use of scarce water supplies (Stavins, 1983; Howe, 1997), and policies that facilitate the restructuring of electricity generation and transmission. The western United States has long been plagued by inefficient use and allocation of its scarce water supplies, largely because users do not have incentives to take actions consistent with economic and environmental values. Economists have noted that federal and state water policies aggravate rather than improve these problems (Anderson, 1983; Frederick, 1986; El-Ashry and Gibbons, 1986; Wahl, 1989). The disparity in water prices over short geographic distances indicates that markets could play a role in solving increasing urban demands for water without the need for new, environmentally-disruptive dams and reservoirs. Reforms have allowed markets in water rights to develop and voluntary exchanges have developed in several states. For example, an agreement was reached to transfer 100,000 acre-feet of water per year from the farmers of the Imperial Irrigation District in southern California to the Metropolitan Water District in the Los Angeles area. Transactions have emerged elsewhere in California, and in Colorado, New Mexico, Arizona, Nevada, and Utah (MacDonnell, 1990).
Since well-functioning markets depend, in part, on the existence of well-informed producers and consumers, information programs can help foster market-oriented solutions to environmental problems (Table 8). These programs have been of two types. Product labeling requirements have been implemented to improve information sets available to consumers. For example, the U.S. Energy Policy and Conservation Act of 1975 specifies that certain appliances and equipment carry labels with information on products’ energy efficiency and estimated energy costs (U.S. Congress, Office of Technology Assessment, 1992). More recently, EPA and the U.S. Department of Energy developed the Energy Star program, in which energy efficient products can display an Energy Star label. And since 1976, the Department of Energy has provided no-cost energy assessments to small and medium-sized manufacturers through its university-based Industrial Assessment Centers (IAC) program. There has been relatively little analysis of the efficacy of such programs, but limited empirical (econometric) evidence suggests that energy-efficiency product labeling has had significant impacts on efficiency improvements, essentially by making consumers and therefore producers more sensitive to energy price changes (Newell, Jaffe, and Stavins, 1999). Also, about half of the projects recommended by assessment teams in the IAC program were subsequently adopted, with firms applying a one to two-year payback period (or about a 50 to 100 percent hurdle rate) to the decisions (Anderson and Newell, 2004).

Another set of information programs has involved reporting requirements. A prominent example is the U.S. Toxics Release Inventory (TRI), established in 1986, which requires firms to make available to the public information on use, storage, and release of...
specific hazardous chemicals. Such information reporting may increase public awareness of firms’ actions, and consequent public scrutiny may encourage firms to alter their behavior, although the evidence on outcomes is mixed (U.S. General Accounting Office, 1992; Hamilton, 1995; Konar and Cohen, 1997; Ananathanarayanan, 1998; Hamilton and Viscusi, 1999).

3.1.2.5. Government subsidy reduction

Government subsidy reduction constitutes another category of market-based instruments. Subsidies are the mirror image of taxes and, in theory, can provide incentives to address environmental problems. Although subsidies can advance environmental quality (see, for example, Jaffe and Stavins, 1995), it is also true that subsidies, in general, have important disadvantages relative to taxes (Dewees and Sims, 1976; Baumol and Oates, 1988). Because subsidies increase profits in an industry, they encourage entry, and can thereby increase industry size and pollution output (Mestelman, 1982; Kohn, 1985).

In practice, rather than internalizing externalities, many subsidies promote economically inefficient and environmentally unsound practices. In such cases, reducing subsidies can increase efficiency and improve environmental quality. For example, because of concerns about global climate change, increased attention has been given to federal subsidies and other programs that promote the use of fossil fuels. An EPA study indicates that eliminating these subsidies would have a significant effect on reducing carbon dioxide (CO₂) emissions (Shelby et al., 1997). The Federal government is involved in the energy sector through the tax system and through a range of individual agency programs. One study indicates that these activities together cost the government $17 billion annually (Koplow, 1993). A substantial share of these U.S. subsidies and programs were enacted during the “oil crises” to encourage the development of domestic energy sources and reduce reliance on imported petroleum. They favor energy supply over energy efficiency. Although there is an economic argument for government policies that encourage new technologies that have particularly high risk or long term payoffs, mature and conventional technologies currently receive nearly 90 percent of the subsidies.

3.1.2.6. Liability rules

Liability rules have been most frequently employed for acute hazards, particularly for toxic waste sites and for the spill of hazardous materials (Menell, 1991). One important example is the Comprehensive Environmental Response, Compensation, and Liability Act of 1980, which established retroactive liability for companies that are found responsible for the existence of a site requiring clean up.

93 The Koplow (1993) study claims that end-use efficiency receives $1 from a wide variety of implicit and explicit federal subsidies for every $35 received by energy supply.

94 On the other hand, federal user charges and insurance premium taxes include significant levies on fossil fuels, and federal tax differentiation has tended to favor renewable energy sources and non-conventional fossil fuels.
Governments can collect cleanup costs and damages from waste producers, waste transporters, handlers, and current and past owners and operators of a site. Similarly, the Oil Pollution Act makes firms liable for cleanup costs, natural resource damages, and third party damages caused by oil spills onto surface waters; and the Clean Water Act makes responsible parties liable for cleanup costs for spills of hazardous substances.

In an ex post regulatory scheme, private polluters can be held liable for damages to remedy the harms they cause to an affected individual or group. In theory, the full costs of their polluting activities will thus be internalized, and polluters will reduce the expected harm of their activity up to the point at which further reductions become more costly than the expected liability they face.

The effectiveness of liability rules depends in part on the ability of victims of pollution to bring actions to recover damages. There are five potential problems. First, environmental harms may be widely dispersed, and so the expected payoff may not justify the cost to an individual victim of bringing a lawsuit (Cropper and Oates, 1992). This collective-action problem can partially be addressed by permitting individuals to bring class actions on behalf of all those harmed by polluters. Second, frequently there are many sources of a given pollutant, and hence the aggrieved party (or parties) may not be able to identify the actual source of the damages. Third, many pollution harms have long latency periods, meaning that by the time the harm has manifested itself, actions are barred because of statutes of limitations. In some jurisdictions, however, such statutes begin to run only with the discovery of the harm, not the imposition of the risk. Fourth, many environmental impacts, such as induced disease, are stochastic by nature, that is, environmental exposure increases the probability of morbidity or mortality. In such cases, it is difficult or impossible to determine with certainty the source of environmental harm. Evidentiary rules that require “a preponderance of the evidence” showing that the plaintiff caused the defendant’s harm would not allow recovery under these circumstances. Fifth, a polluter may not have sufficient solvency to pay a large damage award, and the difference between the polluter’s total solvency and the full damages will be externalized onto the public.

Nevertheless, liability rules have a central role to play in environmental regulation, because other regulatory tools give rise to their own sets of problems. There are important choices that need to be made in designing liability rules, however. Should polluters

---

95 For economic analyses of the Superfund program, see, for example: Hamilton (1993); Gupta, Van Houtven, and Cropper (1996); and Hamilton and Viscusi (1999).

96 Ex post and ex ante legal regimes transmit different incentives to private actors. Shavell (1984) argued that the choice between the two regimes should be considered in light of four factors. First, a liability regime might be preferable if private parties have better information than a regulating authority regarding the risks of productive activities. Second, the greater the likelihood that a private party will not be able to pay fully for a harm, the more attractive is a regulatory regime. Third, the greater the chance that private parties will not face the threat of a lawsuit, the more should regulation be favored. Fourth, the administrative costs associated with the two regimes generally weigh in favor of a liability scheme. Also see: Kolstad, Ulen, and Johnson (1990).

97 This section draws upon Kornhauser and Revesz (2000).

98 Further, a liability scheme may give private actors an incentive to shed their solvency (through dividends to their shareholders, for example) in order to avoid paying large awards (Kornhauser and Revesz, 1990).
be held jointly and severally liable for the harm they cause? Non-jointly liable? Is a negligence rule preferable to a strict liability rule, or vice-versa? Moreover, which parties should be held liable? Polluters? Site owners? Alternative liability regimes transmit different sets of incentives to private actors and can have dramatically different effects.

3.1.2.6.1. Joint and several versus non-joint liability

When a plaintiff’s injury results from the actions of multiple parties, the choice between joint and several and non-joint liability arises. Under joint and several liability, the plaintiff can recover the full amount of damages from any one of the defendants who share responsibility for the damage. Under a system of non-joint liability, a plaintiff can only recover from a defendant the share of damages attributable to that defendant.

Several choices must be made with respect to any joint and several liability regime. First, joint and several liability regimes may or may not allow for contribution, whereby a defendant that has paid a disproportionately large share of a particular damage award will be compensated by parties that have paid disproportionately small shares of that award. Second, contribution shares can be determined either by reference to comparative fault or on a pro rata basis. Third, in the event that a plaintiff settles with one defendant, the regime must specify how much the total damage award against the remaining defendants ought to be reduced (“set-off”). Under a “pro tanto set-off rule,” the plaintiff’s claim against the non-settling defendant is reduced by the amount of the settlement. In contrast, under an apportioned or proportional share set-off rule, the plaintiff’s claim against the non-settling defendant is reduced by the share of the liability attributable to the settling defendant. Fourth, when one defendant settles and another defendant litigates and loses, the regime must specify whether, under the pro tanto set-off rule, the settling defendant is protected against contribution actions. Fifth, the legal regime must also indicate whether settling defendants can bring actions for contribution against defendants who settle for less than their share of liability.

Sixth, joint-and-several regimes sometimes protect non-settling defendants from a plaintiff’s inadequately low settlements with other defendants through a “good faith” hearing on the settlement’s adequacy. And seventh, the regime must specify whether a sued defendant can join a third-party defendant that the plaintiff has declined to name. These choices among rules can have significant impacts on deterrence (Kornhauser and Revesz, 1989, 1990), as well as on the likelihood of inducing settlements.99

3.1.2.6.2. Liability extension

On whom is liability imposed? Assume that there are two groups of actors: waste generators and disposal site owners. One or both could potentially be held liable for problems associated with the disposal of waste. What liability scheme would be preferable on the grounds of efficiency and deterrence?

A legal regime might impose liability solely on the owner of a hazardous waste site and refuse to extend liability to the generators of that waste. Site owners will, under this

---

99 The impact of the possibility of settlement on the choice-of-regime analysis is analyzed in Kornhauser and Revesz (1994a, 1994b).
regime, bear the full costs of the hazardous waste they receive. In a competitive market, disposal site owners will tend to charge generators the marginal cost of disposal. Site owners will have incentives to accept only waste that can properly be disposed of, based on geologic conditions of the site, possible interactions between different types of waste, and other relevant factors. To reduce their disposal costs, generators may seek to reduce the quantity of hazardous waste they produce, and will send their hazardous waste to sites that can process the wastes most effectively—and hence to the sites where wastes will cause the least harm.

While theoretically sound, such full internalization of the site owner’s costs is unlikely in practice. Cleanup costs for hazardous waste sites can be extraordinarily large, and site owners will likely not be sufficiently solvent to pay total cleanup costs in the event of a high-cost problem. If the probability of such an event occurring is not zero, if the site owner’s solvency is less than the full costs associated with that event, and if the site owner does not fully insure against the risk, then the site owner will not bear the full cost, and will charge a price that will not reflect the full cost of remediation (Shavell, 1987; Pitchford, 1995).

There are two possible solutions to the problem of insolvency (Shavell, 2005). First, polluters could be asked to post a bond equal to possible remediation costs. Given the large costs of environmental clean-up, however, such bonds may drive potential polluters out of the market. Second, polluters can be required to carry insurance for potential liabilities. Because insurance companies’ monitoring costs are likely to be high, however, they will only be able to issue insurance based on easily observable factors unrelated to whether the polluter is taking due care to reduce its pollution. Hence, the polluter’s premiums will not be reduced if it takes due care, and a significant moral hazard problem arises. Insurance may therefore be unavailable in the environmental context (Abraham, 1988). Moreover, minimum asset requirements could have socially undesirable effects by banning from the activity actors that derive benefits that are higher than the harms they imposed, even in light of their reduced incentives to take care caused by their limited solvency (Shavell, 2005).

As an alternative, liability could be extended only to the generators of hazardous waste, so that the generators bear the full cleanup costs associated with their waste production. In this case, generators could achieve efficient disposal costs in one of two ways. First, generators could shift liability onto site owners by offering them a payment in exchange for an indemnification agreement. This solution is, as discussed above, hampered by the problem of insolvency. Alternatively, to coordinate efficiently and keep their liability to a minimum, generators could contract among themselves to dispose of specified waste at specified locations. There are two difficulties with this approach. First, transaction costs are likely to be prohibitively large, and generators are therefore

---

100 The average cleanup cost for a site on the Superfund National Priorities List is $30 million, with many exceeding $100 million.

101 A voluntary insurance program will prove inadequate because low-solvency polluters will have no incentive to purchase insurance for a cost they will never bear.
likely to act in a non-cooperative manner (Kornhauser and Revesz, 1994a). Moreover, whether generators are subject to either joint and several or non-joint liability, a strict liability regime applied to a group of generators produces under-deterrence (Kornhauser and Revesz, 1989, 1995).

The second problem relates, again, to solvency. Low-solvency generators have less incentive than high-solvency generators to produce an optimal amount of waste. One would expect that high-solvency generators will thus refuse to contract with their low-solvency counterparts, particularly under joint and several liability regimes, where high-solvency generators may be held liable for the full amount of damages caused by low-solvency generators. This distortion of the contracting patterns can reduce welfare (Boyd and Ingberman, 1997). A non-joint liability regime produces the same result if the damages at the site are allocated among generators proportionally to the amount of waste dumped, and the damage function is convex. In that scenario, one generator’s decision to dump more than the optimal amount results in higher liability for the other generators (Kornhauser and Revesz, 1989).

Finally, given that extending liability solely to site owners or solely to generators results in inefficiencies, liability regimes that target a larger set of parties can achieve two goals. First, such a regime can transmit proper incentives to a larger group of actors. Thus, even if a site owner should become insolvent, generators will have an incentive to continue monitoring. Second, dividing liability between a number of actors can decrease the likelihood of insolvency.

3.1.3. Cross-cutting issues

Three cross-cutting issues stand out in the normative analysis of environmental policy instrument choice: the implications of uncertainty; effects on technological change; and distributional considerations.

3.1.3.1. Implications of uncertainty for instrument choice

The dual task facing policymakers of choosing environmental goals and selecting policy instruments to achieve those goals must be carried out in the presence of the significant uncertainty that affects the benefits and the costs of environmental protection. Since Weitzman’s (1974) classic paper on “Prices vs. Quantities,” it has been generally acknowledged that benefit uncertainty on its own has no effect on the identity of the efficient control instrument, but that cost uncertainty can have significant effects, depending upon the relative slopes of the marginal benefit (damage) and marginal cost functions. In particular, if uncertainty about marginal abatement costs is significant, and if marginal abatement costs are flat relative to marginal benefits, then a quantity instrument is more efficient than a price instrument.102

We rarely encounter situations in which there is exclusively either benefit uncertainty or cost uncertainty. On the contrary, in the environmental arena, we typically find that

102 For an early empirical application in the environmental realm, see: Kolstad (1986).
the two are present simultaneously, and more often than not, it is benefit uncertainty that is of substantially greater magnitude. When marginal benefits are positively correlated with marginal costs (which, it turns out, is not uncommon), then there is an additional argument in favor of the relative efficiency of quantity instruments (Stavins, 1996). On the other hand, the regulation of stock pollutants will often favor price instruments, because the marginal benefit function—linked with the stock of pollution—will tend to be relatively flat, compared with the marginal cost function—linked with the flow of pollution (Newell and Pizer, 2003b).

In theory, there would be considerable efficiency advantages in the presence of uncertainty of hybrid systems—for example, quotas combined with taxes—or non-linear taxes (Roberts and Spence, 1976; Weitzman, 1978; Kaplow and Shavell, 2002b; Pizer, 2002), but such systems have not been adopted.

3.1.3.2. Effects of instrument choice on technological change Environmental policy interventions foster constraints and incentives that affect the process of technological change (Kneese and Schulze, 1975; Orr, 1976). To be dynamically cost-effective, instruments need to foster rather than inhibit technological invention, innovation, and diffusion (Kemp and Soete, 1990). Both command-and-control and market-based instruments have the potential for forcing or inducing technological change, by requiring firms to alter their behavior. Technology and performance standards can be used to stimulate innovation by setting ambitious targets, beyond the reach of current technologies. But it is impossible to know whether a given target will be feasible or not, and so such policies run substantial risk of failure (Freeman and Haveman, 1972). Technology standards are particularly problematic, because they tend to freeze the development of technologies that might otherwise result in greater levels of control.

Much of the economic research on technological invention and innovation (commercialization) has focused on incentives for firm-level decisions to incur costs of research and development in the face of uncertain outcomes. The earliest relevant work was by Magat (1978, 1979), who compared taxes, subsidies, permits, effluent standards, and technology standards, and showed that all but technology standards would induce innovation biased toward emissions reduction. More recent theoretical attempts to rank policy instruments according to their innovation-stimulating effects (Fischer, Parry, and Pizer, 2003) conclude that an unambiguous rating of instruments is not possible. The ranking of instruments depends on the innovator’s ability to appropriate spillover benefits of new technologies to other firms, the costs of innovation, environmental benefit functions, and the number of firms producing emissions (Carraro and Soubeyran, 1996; Katsoulacos and Xepapadeas, 1996; Montero, 2002).

Turning to technological diffusion (adoption), several theoretical studies have found that the incentive for the adoption of new technologies is greater under market-based

---

103 In addition to the efficiency advantages of non-linear taxes, they also have the attribute of reducing the total (although not the marginal) tax burden of the regulated sector, relative to an ordinary linear tax, which is potentially important in a political economy context.
instruments than under direct regulation (Zerbe, 1970; Downing and White, 1986; Milliman and Prince, 1989), but theoretical comparisons among market-based instruments have produced only limited agreement (Milliman and Prince, 1989, 1992; Marin, 1991; Jung, Krutilla, and Boyd, 1996; Parry, 1998; Denicolò, 1999; Keohane, 1999). One overall result seems to be that auctioned permits are inferior in their diffusion incentives to emissions tax systems (but both are superior to command-and-control instruments).104

Closely related to the effects of instrument choice on technological change are the effects of vintage-differentiated regulation on the rate of capital turnover, and thereby on pollution abatement costs and environmental performance. Such vintage-differentiated regulation is a common feature of many environmental and other regulatory policies in the United States, wherein the standard for regulated units is fixed in terms of their date of entry, with later vintages facing more stringent regulation. In the most common application, commonly referred to as “grandfathering,” units produced prior to a specific date are exempted from a new regulation or face less stringent requirements.

While this approach has long appealed to many participants in the policy community, economists have frequently noted that vintage-differentiated regulations can be expected—on the basis of standard investment theory—to retard turnover in the capital stock, and thereby to reduce the cost-effectiveness of regulation, compared with equivalent undifferentiated regulations. Furthermore, under some conditions the result can be higher levels of pollutant emissions than would occur in the absence of regulation. Such economic and environmental consequences are not only predictions from theory (Gruenspecht, 1981; Maloney and Brady, 1988); both types of consequences have been validated empirically in the context of specific regulations (Hartman, Bazdogan, and Nadkarni, 1979; Gruenspecht, 1982; Nelson, Tietenberg, and Donihue, 1993).

3.1.3.3. Distributional considerations Alternative policy instruments can have significantly different impacts on the distribution of benefits and costs. First with regard to benefits, taxes or tradeable permits can lead to localized “hot spots” with relatively high levels of ambient pollution. This is a significant distributional issue, and it can also become an efficiency issue if damages are non-linearly related to pollutant concentrations (Mendelsohn, 1986). The problem can be addressed, in theory, through the use of “ambient permits”105 or through charge systems that are keyed to changes in ambient conditions at specified locations (Revesz, 1996), or through trading schemes that are simply constrained by the requirement that ambient standards not be violated.106 Despite the theoretical literature on ambient systems going back to Montgomery (1972),

104 For a detailed review of analyses of the effects of instrument choice on technological innovation and diffusion, see Jaffe, Newell, and Stavins (2002).
105 Ambient permits entitle the owner to increase the concentration at a certain receptor site by a specified amount, rather than permitting some quantity of emissions.
106 In theory, the locus of regulation can range from input levels (for example, through the use of a permit linked to the carbon content of fossil fuels), to emissions, to ambient concentrations, to exposure levels, to—ultimately—risk levels.
they have never been implemented,\textsuperscript{107} with the partial exception of a two-zone trading system under Los Angeles’ RECLAIM program.\textsuperscript{108}

Turning to the cost side, taxes and auctioned tradeable permits can raise revenue for the government.\textsuperscript{109} Revenue recycling (that is, using tax or permit revenues to reduce other, distortionary taxes) can significantly lower the costs of pollution control, relative to what the costs would be without such recycling (Jorgenson and Wilcoxen, 1994; Goulder, 1995b). It has been suggested by some that all of the abatement costs associated with a pollution tax can be eliminated through revenue recycling (Repetto et al., 1992), but environmental taxes can exacerbate distortions associated with remaining taxes on investment or labor, and research indicates that these distortions are at least as great as those from labor taxes (Bovenberg and de Mooji, 1994; Goulder, 1995a; Parry, 1995; Bovenberg and Goulder, 1996; Goulder, Perry, and Burtraw, 1997; Goulder et al., 1999).

Although distribution affects social welfare, an important strand of the theoretical literature suggests that distribution should not matter in choosing policy instruments. Hylland and Zeckhauser (1979) argue that given optimal taxation achieved by benefit-offsetting tax adjustments, maximizing net benefits should be the sole criterion for policy choice.\textsuperscript{110} Despite its limitations, the income tax system is, in theory, best suited for redistributing income, with attempts at redistribution through other means causing inefficiencies that are at least as great as those encountered with income taxes. Moreover, Kaplow and Shavell (2001) demonstrate that any policy assessment that accords importance to non-utility criteria (including societal concerns for distribution) violates the Pareto principle, suggesting that environmental issues should be addressed ideally through a pair of policies: an efficient environmental policy instrument chosen solely on the basis of maximizing net benefits, and an income tax adjustment to offset possible undesirable distributional impacts.\textsuperscript{111}

Political economy considerations may run counter to such theoretical arguments, since it is difficult to combine every environmental policy rule with a change in the income tax system. It may also be politically infeasible to adopt environmental policies that do not themselves address distributional concerns. Indeed, Arrow et al. (1996b) ar-

\textsuperscript{107} Such systems can be difficult to implement. If there are many significant receptor sites, the implementation of tradeable permits will require separate markets for each type of permit. For a review of ambient permit approaches, see Tietenberg (1995).

\textsuperscript{108} In the case of RECLAIM, empirical analysis indicated that a substantial share of the relevant heterogeneity in concentrations would be captured by employing just two zones (Johnson and Pekelney, 1996).

\textsuperscript{109} While an allocation of permits made through sale or auction will have similar distributional consequences to a tax, a revenue-neutral emissions tax, in which revenues are refunded to regulated firms (but not in proportion to their emissions levels), can resemble—in distributional terms—a comparable tradeable permit program in which the permits are allocated without charge.

\textsuperscript{110} Shavell (1981) and Kaplow and Shavell (1994) extended this result to legal rulemaking.

gue that the Kaldor–Hicks criterion should be considered as neither a necessary nor a sufficient condition for public policy.\(^{112}\)

A policy’s political feasibility is influenced strongly by its distributional implications. Auctioned permit systems or effluent charges can be more costly than comparable command-and-control instruments from the perspective of regulated firms. Tradeable permit systems, on the other hand, have the important attribute that in the absence of decreasing marginal transactions costs (essentially volume discounts), the equilibrium allocation and hence aggregate abatement costs of a tradeable permit system are independent of initial allocations (Stavins, 1995). Hence, the allocation decision can be left to politicians, with limited normative concerns about the potential effects of the chosen allocation on overall cost-effectiveness. In other words, cost-effectiveness or efficiency can be achieved, while distributional equity is simultaneously addressed with the same policy instrument.\(^{113}\)

### 3.1.4. Normative lessons

Although there has been considerable experience in the United States with market-based instruments for environmental protection, this relatively new set of policy approaches has not replaced nor come anywhere close to replacing conventional, command-and-control policies. When and where these approaches have been used in their purest form and with some success, they have not always performed as anticipated. We review briefly the normative lessons that can be learned from research and experience.

#### 3.1.4.1. Design and implementation

The performance to date of market-based instruments for environmental protection provides evidence that these approaches can achieve major cost savings while accomplishing their environmental objectives. The performance of these systems also offers lessons about the importance of flexibility, simplicity, and the capabilities of the private sector.

In regard to flexibility, allowing flexible timing and intertemporal trading of permits—that is, banking allowances for future use—played a very important role in the SO\(_2\) allowance trading program’s performance (Ellerman et al., 1997), much as it did in the U.S. lead rights trading program a decade earlier (Kerr and Maré, 1997).\(^{114}\)

One of the most significant benefits of using market-based instruments may simply

---

\(^{112}\) See the discussion of the Kaldor–Hicks criterion in section 2.1.1.

\(^{113}\) This is one of the reasons why an international tradeable permit mechanism has been considered to be particularly attractive for addressing global climate change. Allocation mechanisms can be developed that address equity concerns of developing countries, and thus increase the political base for support, without jeopardizing the overall cost-effectiveness of the system. See, for example, Frankel (1999). It should be recognized, however, that in practice tradeable permits have typically not been allocated to achieve goals of distributional equity per se, but to achieve political feasibility (Joskow and Schmalensee, 1998).

\(^{114}\) In theory, a fully cost-effective permit trading program must allow for both banking and borrowing (Rubin, 1996; Kling and Rubin, 1997).
be that technology standards are thereby avoided. Less flexible systems would not have led to the technological change that may have been induced by market-based instruments (Burtraw, 1996; Ellerman and Montero, 1998; Bohi and Burtraw, 1997; Keohane, 2001), nor the induced process innovations that have resulted (Doucet and Strauss, 1994).

In regard to simplicity, transparent formulae—whether for permit allocation or tax computation—are difficult to contest or manipulate. Requiring prior government approval of individual trades may increase uncertainty and transaction costs, thereby discouraging trading; these negative effects should be balanced against any anticipated benefits due to requiring prior government approval. Such requirements hampered EPA’s Emissions Trading Program in the 1970s, while the lack of such requirements was an important factor in the success of lead trading (Hahn and Hester, 1989). In the case of SO2 trading, the absence of requirements for prior approval reduced uncertainty for utilities and administrative costs for government, and contributed to low transactions costs (Rico, 1995).

One potentially important cause of the mixed performance of implemented market-based instruments is that many firms are simply not well equipped to make the decisions necessary to fully utilize these instruments. The focus of environmental, health, and safety departments in private firms has been primarily on problem avoidance and risk management, rather than on the creation of opportunities made possible by market-based instruments. Since market-based instruments have been used on a limited basis only, and firms are not certain that these instruments will be a lasting component on the regulatory landscape, it is not surprising that most companies have not reorganized their internal structure to fully exploit the cost savings these instruments offer (Reinhardt, 2000). Rather, most firms continue to have organizations that are experienced in minimizing the costs of complying with command-and-control regulations, not in making the strategic decisions allowed by market-based instruments.115

3.1.4.2. Identifying new applications Market-based policy instruments are considered today for nearly every environmental problem that is raised, ranging from endangered species preservation to global climate change.116 Where the cost of abating pollution differs widely among sources, a market-based system is likely to have greater gains, relative to conventional, command-and-control regulations (Newell and Stavins, 2003). For example, it was clear early on that SO2 abatement cost heterogeneity was great, because of differences in ages of plants and their proximity to sources of low-sulfur

---

115 There are, of course, exceptions. See: Hockenstein, Stavins, and Whitehead (1997). There is anecdotal evidence which may suggest that the existence of tradeable permit programs is changing the way firms evaluate environmental risk (Hartridge, 2003; Tietenberg, 2003).

coal. But where abatement costs are more uniform across sources, the political costs of enacting an allowance trading approach are less likely to be justifiable.

The choice of market-based instrument should depend on the characteristics of the pollutant, the degree of uncertainty, expected changes in the economic environment, ability to induce technological change, potential transactions costs, and other significant interacting factors. Finally, considerations of political feasibility point to the wisdom (more likely success) of proposing market-based instruments when they can be used to facilitate a cost-effective, aggregate emissions reduction (as in the case of the SO2 allowance trading program in 1990), as opposed to a cost-effective reallocation of the status quo burden.

3.2. Positive issues and analysis

A set of positive political economy questions are raised by the increasing use of market-based instruments for environmental protection. First, why was there so little use of market-based instruments in the United States, relative to command-and-control instruments, over the 30-year period of major environmental regulation that began in 1970, despite the apparent advantages these instruments offer? Second, when market-based instruments have been adopted, why has there been such great reliance on tradeable permits allocated without charge, despite the availability of a much broader set of incentive-based instruments? Third, why has the political attention given to market-based environmental policy instruments increased dramatically in recent years?

To examine these questions, we employ a “political market” metaphor (Keohane, Revesz, and Stavins, 1998). The commodity being supplied is legislators’ support for a given policy instrument, and the currency is resources that can be used for re-election—contributions, endorsements, votes, and other forms of support. Demand for legislative outcomes comes from interest groups, including environmental advocacy organizations, private firms, industry associations, organized labor, and consumers. Ultimately, the choice of environmental policy instrument is determined by the equilibrium between demand by interest groups and supply by legislators and regulators.

3.2.1. Historical dominance of command-and-control

On the regulatory demand side, affected firms and their trade associations prefer instruments that have lower aggregate costs for their industry, or that increase aggregate profits by creating rents or barriers to entry. An individual firm may actually prefer regulation to the status quo if that regulation gives the firm a competitive advantage over rivals. Command-and-control standards have the potential to

---

117 This section of the chapter draws upon: Keohane, Revesz, and Stavins (1998); Hahn and Stavins (1991); and Stavins (1998).

118 There are other possible explanations for firms’ preferences, including the possibility that existing agents tend to support the status quo for fear that their expertise will be devalued under new regimes. There is also
generate rents for existing firms in an industry, increasing aggregate profits. Regulations that establish long-term barriers to entry can sustain these profits indefinitely and will be strongly preferred by industry groups (Buchanan and Tullock, 1975; Maloney and McCormick, 1982). Command-and-control standards are inevitably set up with extensive input from industry, and frequently contain more stringent requirements for new sources and other effective barriers to entry (Stigler, 1971; Rasmusen and Zupan, 1991). Firms also tend to favor command-and-control regulation or grandfathered permits over pollution taxes or auctioned permits because they shift less of the distributional burden onto private industry (Arnold, 1995; Crandall, 1983; Hahn and Noll, 1990). Auctioned permits and pollution taxes require firms to pay not only abatement costs, but also regulatory costs in the form of permit purchases or tax payments.

Environmental advocacy groups are also on the demand side of the market for legislative support. Such groups seek to maximize their utility, which depends on both their organizational well-being and the level of environmental quality. For a long time, nearly all environmental advocacy groups were actively hostile towards market-based instruments. One reason was philosophical: environmentalists frequently perceived pollution taxes and tradeable permits as “licenses to pollute.” Although such ethical objections to the use of market-based environmental strategies have greatly diminished, they have not disappeared completely (Sandel, 1997). A second concern was that damages from pollution were difficult or impossible to quantify and monetize, and thus could not be summed up in a marginal damage function or captured by a Pigovian tax rate (Kelman, 1981a). Third, environmental organizations have opposed market-based schemes for strategic reasons, particularly the fear that permit levels and tax rates—once implemented—would be more difficult to tighten over time than command-and-control standards. For example, if permits are given the status of “property rights,” then any subsequent attempt by government to reduce pollution levels further could meet with demands for compensation. Finally, environmental organizations have objected to decentralized instruments on the technical grounds that even if emission taxes or tradeable permits reduce overall levels of emissions, they can—in theory—lead to localized “hot spots” with relatively high levels of ambient pollution.

The final influential group demanding support from legislators is organized labor. Labor groups can be expected to seek protection for jobs, and they may oppose instruments that are likely to lead to plant closings or major industrial dislocations. For the possibility that market-based instruments were opposed simply because they were not well understood (Kelman, 1981b).

119 Sandel (1997) argues that emissions trading will foster “immoral” behavior by giving firms a “license to pollute,” despite the fact that pollution taxes and tradeable permits create incentives for firms to decrease pollution. See Shavell et al. (1997) for replies to Sandel’s arguments. For a broader examination of the ethical limitations of markets in other arenas, see Sandel (1998).

120 This concern was alleviated in the SO2 provisions of the Clean Air Act Amendments of 1990 by an explicit statutory provision that permits do not represent property rights.
example, labor might oppose a tradeable permit scheme in which firms would tend to close factories in heavily polluted areas, sell permits, and relocate to less polluted areas where permits are cheaper (Hahn and Noll, 1990). In the case of restrictions on clean air, organized labor has taken the side of the United Mine Workers, whose members are heavily concentrated in eastern mines that produce higher-sulfur coal, and had therefore opposed pollution-control measures that would increase incentives for using low-sulfur coal from the largely non-unionized (and less labor-intensive) mines in Wyoming’s and Montana’s Powder River Basin. In the 1977 debates over amendments to the Clean Air Act, organized labor fought to include a command-and-control standard that effectively required scrubbing, thereby seeking to discourage switching to cleaner western coal (Ackerman and Hassler, 1981). Likewise, the United Mine Workers opposed the SO2 allowance trading system in 1990, because of a fear that it would encourage a shift to western low-sulfur coal from non-unionized mines.

Turning to the supply side of environmental regulation, legislators may be thought of as providing support as a function of the opportunity cost of supporting a given instrument, the psychological cost associated with their ideological preferences, and the losses or gains of constituency support as a result of an action (Keohane, Revesz, and Stavins, 1998). Legislators have had a number of reasons to find command-and-control standards attractive. First, many legislators and their staffs are trained in law, which may predispose them to favor conventional regulatory approaches and lead to a status quo bias in favor of command-and-control approaches (Kneese and Schulze, 1975). Second, standards tend to help hide the costs of pollution control (McCubbins and Sullivan, 1984; Hahn, 1987), while market-based instruments generally impose those costs more directly. Third, standards offer greater opportunities for symbolic politics, because strict standards—strong statements of support for environmental protection—can readily be combined with less visible exemptions or with lax enforcement measures.

Fourth, if politicians are risk averse, they will prefer instruments that involve more certain effects.121 The flexibility inherent in market-based instruments creates uncertainty about distributional impacts. Typically, legislators in a representative democracy are more concerned with the geographic distribution of costs and benefits than with comparisons of total benefits and costs. Hence, aggregate cost-effectiveness—the major advantage of market-based instruments—is less likely to play a significant role in the legislative calculus than whether a politician is getting a good deal for his or her constituents (Shepsle and Weingast, 1984).

Finally, legislators are wary of enacting programs that are likely to be undermined by bureaucrats in their implementation. And bureaucrats are less likely to undermine legislative decisions if their own preferences over policy instruments are accommodated. Bureaucratic preferences—at least in the past—were not supportive of market-based

---

121 Legislators are likely to behave as if they are risk averse, even if they are personally risk neutral, if their constituents punish unpredictable policy choices or their reelection probability is nearly unity (McCubbins, Noll, and Weingast, 1989).
3.2.2. Prevalence of tradeable permits allocated without charge

Economic theory suggests that the choice among tradeable permits, pollution taxes, and other market-based instruments should be based upon case-specific factors, but major applications in the United States have nearly always taken the form of tradeable permits allocated without charge, rather than through auctions, despite the apparent economic superiority of the latter mechanism in terms of economic efficiency. Many participants in the policy process have reasons to favor tradeable permits allocated without charge over other market-based instruments.

On the regulatory demand side, existing firms favor tradeable permits allocated without charge because such permits convey rents to firms. Moreover, like stringent command-and-control standards for new sources, but unlike auctioned permits or taxes, permits allocated without charge give rise to entry barriers, since new entrants must purchase permits from existing holders. Thus, the rents conveyed to the private sector by tradeable permits allocated without charge are, in effect, sustainable.

Environmental advocacy groups have generally supported command-and-control approaches, but given the choice between tradeable permits and emission taxes, these groups strongly prefer the former. Environmental advocates have a strong incentive to avoid policy instruments that make the costs of environmental protection highly visible to consumers and voters; and taxes make those costs more explicit than permits. Also, environmental advocates prefer permit schemes because they specify the quantity of pollution reduction that will be achieved, in contrast with the indirect effect of pollution taxes. Overall, some environmental groups have come to endorse the tradeable permits approach because it promises the cost savings of pollution taxes, but without the drawbacks that environmentalists associate with tax instruments.

Tradeable permits allocated without charge are easier for legislators to supply than taxes or auctioned permits, again because the costs imposed on industry are less visible and less burdensome, since no money is exchanged at the time of the initial permit allocation. Also, permits allocated without charge offer a much greater degree of political control over the distributional effects of regulation, facilitating the formation of majority coalitions. Joskow and Schmalensee (1998) examined the political process of

---

122 Subsequently, this same incentive led EPA staff involved in the acid rain program to become strong proponents of trading for a variety of other pollution problems.

123 The EPA does have an annual (revenue-neutral) auction of SO2 allowances, but this represents less than 2 percent of the total allocation (Bailey, 1996). While the EPA auctions may have helped in establishing the market for SO2 allowances, they are a trivial part of the overall program (Joskow, Schmalensee, and Bailey, 1998).
allocating SO₂ allowances in the 1990 amendments, and found that allocating permits on the basis of prior emissions can produce fairly clear winners and losers among firms and states. An auction allows no such political maneuvering.

3.2.3. Increased attention to market-based instruments

Interest in and use of incentive-based instruments has increased at both the Federal and state levels in recent years (Hahn, 2000). Given the historical lack of receptiveness by the political process to market-based approaches to environmental protection, why has there been this rise in the use of these approaches? It would be gratifying to believe that increased understanding of market-based instruments had played a large part in fostering their increased political acceptance, but how important has this really been?

In 1981, Kelman surveyed Congressional staff members, and found that support and opposition to market-based environmental policy instruments was based largely on ideological grounds: Republicans, who supported the concept of economic-incentive approaches, offered as a reason the assertion that “the free market works,” or “less government intervention” is desirable, without any real awareness or understanding of the economic arguments for market-based programs. Likewise, Democratic opposition was largely based upon ideological factors, with little or no apparent understanding of the real advantages or disadvantages of the various instruments (Kelman, 1981b). What would happen if we were to replicate Kelman’s survey today? Our refutable hypothesis is that we would find increased support from Republicans, greatly increased support from Democrats, but insufficient improvements in understanding to explain these changes.¹²⁴ So what else has mattered?

First, one factor has surely been increased pollution control costs, which have led to greater interest from all parties in cost-effective instruments. By the late 1980’s, even political liberals and environmentalists were beginning to question whether conventional regulations could produce further gains in environmental quality. During the previous twenty years, pollution abatement costs had continually increased, as stricter standards moved the private sector up the marginal abatement-cost function. By 1990, U.S. pollution control costs had reached $125 billion annually, nearly a 300% increase in real terms from 1972 levels (U.S. Environmental Protection Agency, 1990; Jaffe et al., 1995). Market-based instruments represent an effective way to reduce aggregate abatement costs.

Second, a factor that became important in the late 1980’s was strong and vocal support from some segments of the environmental community. By supporting tradeable permits for acid rain control, the Environmental Defense Fund (EDF) seized a market niche in the environmental movement, and successfully distinguished itself from

¹²⁴ But there has been some increased understanding of market-based approaches among policy makers. This has partly been due to increased understanding by their staffs, a function—to some degree—of the economics training that is now common in law schools, and of the proliferation of schools of public policy (Hahn and Stavins, 1991).
other groups. Related to this, a third factor was that the SO2 allowance trading program, the leaded gasoline phasedown, and the CFC phaseout were all designed to reduce emissions, not simply to reallocate them cost-effectively among sources. Market-based instruments are most likely to be politically acceptable when proposed to achieve environmental improvements that would not otherwise be feasible (politically or economically).

Fourth, deliberations regarding the SO2 allowance system, the lead system, and CFC trading differed from previous attempts by economists to influence environmental policy in an important way: the separation of ends from means, i.e. the separation of consideration of goals and targets from the policy instruments used to achieve those targets. By accepting the politically identified (and potentially inefficient) goal, the ten-million ton reduction of SO2 emissions, for example, economists were able to focus successfully on the importance of adopting a cost-effective means of achieving that goal. Fifth, acid rain was an unregulated problem until the SO2 allowance trading program of 1990; and the same can be said for leaded gasoline and CFC’s. Hence, there were no existing constituencies—in the private sector, the environmental advocacy community, or government—for the status quo approach, because there was no status quo approach.

Sixth, by the late 1980’s, there had already been a perceptible shift of the political center toward a more favorable view of using markets to solve social problems. The George H. W. Bush Administration, which proposed the SO2 allowance trading program and then championed it through an initially resistant Democratic Congress, was (at least in its first two years) “moderate Republican;” and phrases such as “fiscally responsible environmental protection” and “harnessing market forces to protect the environment” do have the sound of quintessential moderate Republican issues. But, beyond this, support for market-oriented solutions to various social problems had been increasing across the political spectrum for the previous fifteen years, as was evidenced by deliberations on deregulation of the airline, telecommunications, trucking, railroad, and banking industries. Indeed, by the mid-1990s, the concept (or at least the phrase), “market-based environmental policy,” had evolved from being politically problematic to politically attractive.

---

125 When the memberships (and financial resources) of other environmental advocacy groups subsequently declined with the election of the environmentally-friendly Clinton-Gore Administration, EDF continued to prosper and grow (Lowry, 1993). EDF has since renamed itself “Environmental Defense.”

126 The Reagan Administration enthusiastically embraced a market-oriented ideology, but demonstrated little interest in employing actual market-based policies in the environmental area. From the Bush Administration through the Clinton Administration, interest and activity regarding market-based instruments—particularly tradeable permit systems—continued to increase, although the pace of activity in terms of newly implemented programs declined during the Clinton years, when a considerable part of the related focus was on global climate policy (Hahn, Olmstead, and Stavins, 2003).
4. Allocation of responsibility across levels of government

Throughout most of U.S. history, state and local governments have had primary responsibility for health-and-safety regulation, including environmental protection, but since 1970, the Federal government has played an increasingly important role in environmental regulation (Revesz, 2001a). What regulatory advantages or disadvantages does the Federal government have, compared with state governments? What does this suggest about how regulatory responsibility should be allocated?

4.1. Positive review of responsibility of levels of government

Before 1970, Congress largely left environmental regulation to the states. As the modern environmental movement gained political force, however, the Federal government began assembling its regulatory framework. Congress’s first major effort came in 1969 with the passage of the National Environmental Policy Act, which laid out broad environmental goals and required Federal agencies to assess the environmental impacts of their programmatic actions.

A set of laws passed over the subsequent two decades marked the federal government’s new-found commitment to environmental regulation. Three statutes formed the backbone of the federal scheme: the Clean Air Act of 1970, the Clean Water Act of 1972, and the Resource Conservation and Recovery Act of 1976, all of which have been amended numerous times since their adoption. These laws, characterized by their command-and-control approaches to regulation, granted wide discretion to the newly-created Environmental Protection Agency to set tolerance levels for various pollutants. In addition, Congress protected endangered species, set limits on contaminants allowed in drinking water, and created a system of strict joint-and-several liability for parties responsible for abandoned hazardous waste sites.

Federal environmental laws typically (but with important exceptions) establish minimum environmental standards while leaving states free to adopt more stringent standards. Many states have done exactly that. Some have adopted tighter thresholds for automobile emissions; others have created their own “Superfund” programs; and others have implemented their own state-based environmental protection acts (Revesz, 2001a). States also remain free to regulate in areas the federal government has opted not to, such as wetlands preservation or groundwater quality.

130 33 U.S.C. §§1251–1376. The Act was originally titled the Federal Water Pollution Control Act.
133 Safe Drinking Water Act, 42 U.S.C. 300f et seq.
4.2. Normative review of allocation of regulatory responsibility

A rebuttable presumption in favor of decentralized authority over environmental regulation may be posited for three reasons. First, in a large and diverse country, different regions will likely have different environmental preferences. Second, the benefits of environmental protection vary throughout the country. For example, a stringent air quality standard may benefit many people in densely populated areas but only a few elsewhere. Third, the costs of meeting a given standard differ across geographic regions.

Federal intervention in environmental regulation has traditionally been justified by reference to one or more of three perceived pathologies that hamper effective state regulation: the race to the bottom induced by competition for mobile resources; the existence of significant interstate externalities; and the public-choice rationale that environmental groups can more effectively lobby at the Federal level than at the state level. Analysis has cast doubt on the viability of these justifications.135

4.2.1. Competition among political jurisdictions: the race to the bottom

The conventional race-to-the-bottom rationale for federal regulation posits that states, in an effort to induce geographically mobile firms to locate within their jurisdictions, will offer them sub-optimally lax environmental standards in order to benefit from additional jobs and tax revenues. In the absence of Federal regulation, states would therefore systematically under-regulate.

4.2.1.1. Normative assessment of the race-to-the-bottom claim  The theoretical foundation for the view that interstate competition for industry would inevitably lead to sub-optimally lax environmental standards is weak. Indeed, economic analysis of the effects of interstate competition on the choice of environmental standards indicates that rather than a race to the bottom, inter-jurisdictional competition may be expected to lead to the maximization of social welfare, at least under conditions of perfect competition (Oates and Schwab, 1988). In their model, Oates and Schwab posit jurisdictions that compete for mobile capital through the choice of taxes and environmental standards. A higher capital stock benefits residents in the form of higher wages, but hurts them through foregone tax revenues and lower environmental quality. Each jurisdiction makes two policy decisions: it sets a tax rate on capital and an environmental standard. Oates and Schwab show that competitive jurisdictions will set a net tax rate on capital of zero (the rate that exactly covers the cost of public services provided to the capital, such as police and fire protection). In turn, competitive jurisdictions will set an environmental standard that is defined by equating the willingness to pay for an additional unit of environmental quality with the corresponding change in wages. Oates and Schwab show that these choices of tax rates and environmental standards are efficient.

135 This section draws upon Revesz (2001b). Also see Krier (1995).
When the assumption of perfect competition is relaxed, strategic interactions among the states can lead to under-regulation absent federal intervention. But it is likewise plausible that in other instances the reverse would be true: that the strategic interactions among the states would lead to over-regulation absent federal intervention. Accordingly, there is no compelling race-to-the-bottom justification for across-the-board federal minimum standards, the cornerstone of Federal environmental law.

The most extensive analyses of the effects of imperfect competition among states show that either over-regulation or under-regulation can result (Markusen, Morey, and Olewiler, 1993, 1995), depending on the levels of firm-specific costs, plant-specific costs, and transportation costs. Similarly, if a firm has market power enabling it to affect prices, it will be able to extract a sub-optimally lax standard; but if a state has market power, the reverse would be true (Revesz, 1992, 1997b). In summary, just as there are situations in which interstate competition produces environmental under-regulation (Esty and Geradin, 2001), there are other plausible scenarios under which the result is over-regulation.

Moreover, even if states systematically enacted sub-optimally lax environmental standards, Federal environmental regulation would not necessarily improve the situation. If states cannot compete over environmental regulation because it has been federalized, they will compete along other regulatory dimensions, leading to sub-optimally lax standards in other areas, or along the fiscal dimension, leading to the under-provision of other public goods. Thus, the reduction in social welfare implicit in race-to-the-bottom arguments would not be eliminated by federalizing environmental regulation. Rather, the federalization of all regulatory and fiscal decisions would be necessary to solve the problem.136

Several authors have attempted to rehabilitate some version of the race-to-the-bottom justification for Federal regulation. Their arguments, however, rely on confusions of alternative justifications for environmental regulation, such as the presence of inter-jurisdictional externalities or public choice failures (Esty, 1996; Esty and Geradin, 2001), unsupported public-choice rationales that are analytically distinct from the race-to-the-bottom justification (Swire, 1996), weak empirical support (Engel, 1997), or circular notions that Federal environmental regulation serves to reinforce “national evaluative norms” (Sarnoff, 1997). The critics therefore fail to address two core difficulties confronting supporters of Federal environmental regulation (Revesz, 1997b). First, none are able to explain why Federal environmental floors are an appropriate response to races that can lead either to over-regulation or under-regulation. Regulatory ceilings would be, after all, the appropriate Federal response to widespread over-regulation. Second, their arguments for federalizing environmental decision-making prove too much, and tend equally to support the claim that all state fiscal and regulatory decisions should be addressed at the Federal level.

136 Similarly, there is a concern that absent federal regulation, firms could capture rents created by locational advantages that otherwise would accrue to the states. But if environmental regulation is federalized, the rents could be captured with respect to another component of costs. Only complete centralization would address the problem (Engel and Rose-Ackerman, 2001).
4.2.1.2. Positive assessment of the race-to-the-bottom claim

The validity of race-to-the-bottom arguments for federal regulation cannot be resolved on theoretical grounds alone. Empirical analysis is required, and available evidence indicates that the stringency of environmental regulation does not have a statistically significant effect on plant location decisions (Bartik, 1988b, 1989; Friedman, Gerlowski, and Silberman, 1992; Levinson, 1996; McConnell and Schwab, 1991).

More generally, the empirical economic literature on the effects of environmental regulation provides little evidence to support the hypothesis that environmental regulation has a significant adverse effect on economic growth or on other measures of competitiveness (Jaffe et al., 1995). Recent studies have reinforced this conclusion, finding that environmental regulation does not reduce labor demand (Berman and Bui, 2001a), and does not impair productivity (Berman and Bui, 2001b). Such findings are not surprising, given that for all but the most heavily polluting industries, the costs of complying with environmental regulation are a small share of the total costs of production—an average of about 2 percent (Jaffe et al., 1995). It follows that the difference in production costs among jurisdictions with relatively more stringent and relatively less stringent environmental standards is even less. Other regulatory factors—including the level of state taxes, the provision of public services, and the degree of unionization of a state’s labor force—have been shown empirically to exert significant influences on location decisions, just as environmental regulations have not (Bartik, 1988b, 1989; Levinson, 1996). Moreover, there is evidence that large national or multinational firms build their plants to meet the standards of the most stringent jurisdiction in which they have production facilities. Thus, they do not benefit from lower costs of environmental regulation when they operate in jurisdictions with laxer standards (Jaffe et al., 1995).

Of course, even if empirical evidence indicated that firms move from or do not locate in jurisdictions with more stringent environmental standards, this would not necessarily indicate that such a “race-to-the-bottom” was welfare-decreasing. A study of firm mobility measures only what states lose as a result of more stringent environmental standards; it does not assess the corresponding gains that may result from better environmental quality. A state that makes its environmental standards more stringent and thereby loses some economic activity may well increase its social welfare, if the environmental gains are greater than the losses.

4.2.2. Transboundary environmental problems

In contrast to the race-to-the-bottom argument, the presence of interstate externalities provides a potentially sound theoretical argument for Federal regulation. A state that sends pollution to another state can obtain the benefits of the economic activity that generates the pollution, but not suffer the full costs of that activity (Revesz, 1996). Transaction costs—particularly in the case of air pollution—are likely to be sufficiently high to prevent the formation of interstate compacts.\textsuperscript{137}

\textsuperscript{137} It is difficult for such compacts to emerge in the absence of a clearly defined baseline regarding when upwind states have the right to send pollution downwind, and in the absence of generally accepted models
The fact that interstate externalities provide a sensible justification for federal intervention does not mean that existing federal environmental regulations can be justified on these grounds. For some environmental problems, such as the control of drinking water quality, there are virtually no interstate externalities; the effects are almost exclusively local. Even with respect to problems for which interstate externalities exist, the rationale calls only for a response well targeted to the problem, such as a limit on the quantity of pollution that can cross state lines, rather than across-the-board Federal regulation.

In fact, the environmental statutes have been an ineffective response to the problem of interstate externalities. The core of the Clean Air Act, which addresses the type of pollution for which externalities are believed to be most prevalent, consists of a series of Federally prescribed ambient standards and emissions standards.\textsuperscript{138} The federal emission standards do not effectively combat the problem of interstate externalities, because they do not regulate the number of sources within a state or the location of those sources. Similarly, the federal ambient air quality standards are not well targeted to address the problem of interstate externalities, since they require states to restrict pollution that may have only in-state consequences, and states can meet the ambient standards but still export pollution to downwind states (through tall stacks or locations near the interstate border). In fact, a state might meet its ambient standards precisely because it exports a large proportion of its pollution.\textsuperscript{139} Sections 110(a)(2)(D) and 126(b),\textsuperscript{140} enacted in 1977, are the only provisions of the Clean Air Act specifically designed to combat interstate externalities. They create a mechanism by which downwind states can seek to enjoin excessive upwind pollution. During the first two decades of the program, however, no downwind state prevailed on such a claim.

\textit{4.2.3. Public choice and systematic bias}

Advocates for Federal regulation on public choice grounds typically assert that state political processes undervalue the benefits of environmental regulation, or overvalue the corresponding costs. Even if this is true, of course, it does not follow that federalizing environmental law will necessarily provide a solution. Federal regulation is justifiable only if the outcome at the Federal level is socially more desirable, either because there


\textsuperscript{139} The Federal environmental statutes have exacerbated, rather than ameliorated, the problem of interstate externalities. In the context of the Clear Air Act, the Federal ambient standards give states an incentive to encourage sources within their jurisdiction to use taller stacks (or to locate close to downwind borders). Not surprisingly, the use of tall stacks expanded considerably after the passage of the Clean Air Act in 1970, when only two stacks in the United States were higher than 500 feet. By 1985, more than 180 stacks were higher than 500 feet, and 23 were higher than 1,000 feet (Revesz, 1996).

\textsuperscript{140} 42 U.S.C. §§7410(D), 7426 (1994).
is less under-regulation or because any over-regulation leads to smaller social welfare losses. There are several reasons for being skeptical about the soundness of such a claim.

4.2.3.1. Normative foundation for public choice claims The public choice mechanism that makes it possible for Federal regulation to correct for under-regulation at the state level is far from self-evident. For example, Esty (1996) states that “[a]t the centralized level, environmental groups find it easier to reach critical mass and thereby to compete on more equal footing with industrial interests.” He adds that the difficulty of mobilizing the public in many separate jurisdictions is well established. In fact, the logic of collective action may suggest the opposite: given the costs of organizing necessarily larger groups at the Federal level, those groups will likely prove less effective there than at the state level. Aggregating environmental interests on a national level increases the heterogeneity of environmental policy priorities, thereby complicating organizational challenges. The situation is likely to be different for regulated industry groups, which frequently consist of firms with nationwide operations. For such firms, operating at the Federal level poses no additional free-rider problems or loss of homogeneity.

The relevant question is whether the additional problems faced by environmental groups at the Federal level are outweighed by benefits arising from the fact that the clash of interest groups takes place before a single legislature, a single administrative agency, and, in part, as a result of the exclusive venue of the U.S. Court of Appeals for the D.C. Circuit over important environmental statutes, in a single court (Revesz, 1997a). It might be the case that economies of scale of operating at the Federal level would more than outweigh the increased free-rider problems. Such economies of scale might well hold for certain costs associated with effective participation in the regulatory process, such as for hiring a competent scientist. But the structure of other political costs is likely to be quite different. For example, with respect to access to the legislative process, a standard public choice account is that the highest bidder prevails (Peltzman, 1976; Stigler, 1971). Thus, the benefit that a party receives from its expenditures is a function of the expenditures of the other party. Unless costs of this type are small, economies of scale of operating at the federal level are unlikely to outweigh the additional free-rider problems.

Given the standard public choice argument for federal environmental regulation, it is not clear why the problems observed at the state level would not be replicated at the Federal level (Revesz, 1997c). The logic of collective action would suggest that the large number of citizen-breathers, each with a relatively small stake in the outcome of a particular standard-setting proceeding, will be overwhelmed in the political process by concentrated industrial interests with a large stake in the outcome. This problem could occur at the Federal level as well as at the state level.

4.2.3.2. Positive support for public choice pathologies Public choice arguments for federal regulation rest on two empirical claims concerning the nature of state regulatory actions: (1) that states ignored environmental problems before 1970, when the major
environmental statutes began to be enacted; and (2) that states continue to be less concerned about environmental problems than is the Federal government.

First, the view that the states ignored environmental problems before 1970 is simply not correct. Several studies show that during the 1960s, without Federal prodding, states were making considerable strides with respect to the control of air pollution. In particular, the concentrations of important air pollutants were falling at significant rates, and the number of states, counties, and municipalities with regulatory programs to control air pollution was increasing rapidly (Crandall, 1983; Goklany, 1998a, 1998b; Portney, 1990; Stern, 1982). For example, sulfur dioxide concentrations fell by about 11 percent annually between 1964 and 1971, but only by about 5 percent per year in the decade after the Federal government began regulating. Similarly, concentrations of total suspended particulates dropped sharply during the 1960s, but the pace of reduction slowed significantly in the 1970s (Crandall, 1983). The genesis of Federal environmental regulation is consistent with this evidence. The Clean Air Act of 1970 was a response to industry pressure for Federal regulation as a means of discouraging states from setting more stringent (and hence non-uniform) standards (Elliott, Ackerman, and Millian, 1985).

The second view—that states are less concerned about environmental problems than the Federal government—can be countered with reference to many states’ innovative environmental laws which impose (sometimes significant) constraints on in-state firms. In the areas of automobile emission standards, hazardous waste regulation, non-hazardous solid waste regulation, and wetlands protection, states have taken an active role in (effectively) regulating to improve environmental quality, and this involvement increased in the 1990s in the face of the Federal government’s less aggressive action on environmental matters.

Clearly not every state is equally active in the environmental regulatory arena. The citizens of some states may have preferences for laxer environmental regulation than the Federal regulatory level and may therefore not have any reason to adopt voluntarily additional environmental constraints. Indeed, an analysis of Federal representatives’ voting records on issues of environmental concern indicates a strong correlation between support for “pro-environment” bills in Congress and heightened in-state environmental regulatory programs. The existence of significant state regulation calls into question the simplistic public choice claim that environmental groups are less able to lobby effectively at the state level than at the federal level (Revesz, 2001a).

5. Conclusions

The growing use of economic analysis to inform environmental decision making marks increasing acceptance of the usefulness of these tools to help focus and improve regulation. Debates about the normative standing of the Kaldor–Hicks criterion and the challenges inherent in making benefit–cost analysis operational will likely continue.
Nevertheless, economic analysis has assumed a significant position in the regulatory state.

At the same time, despite the arguments made for decades by economists and others, there seems to be no more than limited political support in the United States for much broader use of benefit–cost analysis to assess proposed or existing environmental regulations. In truth, these analytical methods remain on the periphery of policy formulation. In fact, as long as leaders on both sides of the debates in the policy community continue to react on ideological bases to proposals for such “regulatory reform,” the status quo is unlikely to change. Perhaps the significant changes that have taken place over the past twenty years with regard to the means of environmental policy—that is, acceptance of market-based environmental instruments—can provide a model for progress with regards to analysis of the ends—the targets and goals—of public policies in this domain.

Certainly the change has been dramatic. Market-based instruments have moved center stage, and policy debates today look very different from those twenty years ago, when these ideas were routinely characterized as “licenses to pollute” or dismissed as completely impractical. Market-based instruments are now considered seriously for each and every environmental problem that is tackled, ranging from endangered species preservation to regional smog to global climate change. It is reasonable to anticipate that market-based instruments will enjoy increasing acceptance in the years ahead.

Of course, no particular form of government intervention, no individual policy instrument—whether market-based or conventional—and no specific level of government is appropriate for all environmental problems. Which instrument or level of government is best in any given situation depends upon a variety of characteristics of the environmental problem, and the social, political, and economic context in which it is being regulated. There is no policy panacea. But economic instruments are now part of the available policy portfolio, and ultimately that is good news both for environmental protection and economic well-being.

Acknowledgements

Katharine Emans, Nicholas Bagley, and Gernot Wagner provided valuable research assistance, and helpful comments on previous versions were provided by Dallas Burtraw, Daniel Esty, Myrick Freeman, Robert Hahn, Louis Kaplow, Nathaniel Keohane, Charles Kolstad, Peter Menell, Richard Newell, Sheila Olmstead, Mitchell Polinsky, Paul Portney, Steven Shavell, and Kerry Smith. The authors alone are responsible for any remaining errors.

References


Kemp, R., Soete, L. (1990). “Inside the ‘Green Box’: On the Economics of Technological Change and the Environment”. In: Freeman, C., Soete, L. (Eds.), New Explorations in the Economics of Technological Change. Pinter, London.


