ENVIRONMENTAL ECONOMICS

Robert N. Stavins
John F. Kennedy School of Government, Harvard University
and Resources for the Future

Prepared for
Lawrence Blume and Steven Durlauf
General Editors
London: Palgrave Macmillan Ltd.

December 23, 2004

Send Comments to:
Prof. Robert N. Stavins
John F. Kennedy School of Government
Harvard University
79 John F. Kennedy Street
Cambridge, Massachusetts 02138
Phone: 617-495-1820
Fax: 617-496-3783
E-Mail: robert_stavins@harvard.edu
ABSTRACT

This article provides an overview of the economics of environmental policy, including the setting of goals and targets, notably the Kaldor-Hicks criterion and the related method of assessment known as benefit-cost analysis. Also reviewed are the means of environmental policy, that is, the choice of specific policy instruments, featuring an examination of potential criteria for assessing alternative instruments, with focus on cost-effectiveness. The theoretical foundations and experiential highlights of individual instruments are reviewed, including conventional command-and-control mechanisms and market-based instruments.

Keywords: environmental economics, efficiency, cost-effectiveness, benefit-cost analysis, market-based instruments, tradeable permits, pollution taxes

JEL Classification: K320, Q280, Q380, Q480
ENVIRONMENTAL ECONOMICS
(5,799 words; 7,812 words with bibliography)
Robert N. Stavins*

The fundamental theoretical argument for government activity in the environmental realm is that pollution is an externality — an unintended consequence of market decisions, which affects individuals other than the decision maker. Providing incentives for private actors to internalize the full costs of their actions was long thought to be the theoretical solution to the externality problem. 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 response to the Pigovian perspective was provided by Ronald Coase in *The Problem of Social Cost* (1960). Coase demonstrated that in a bilateral bargaining environment with no transaction costs, wealth or income effects, or third-party impacts, two negotiating parties will reach socially desirable agreements, and the overall amount of pollution will be independent of the assignment of property rights. At least some of the specified conditions are unlikely to hold for most environmental problems. Hence, private negotiation will not — in general — fully internalize environmental externalities.

**Criteria for Environmental Policy Evaluation**

More than 100 years ago, Vilfredo Pareto enunciated the well-known normative criterion for judging whether a social change 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. 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 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. 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:

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

---

*Albert Pratt Professor of Business and Government at the John F. Kennedy School of Government, Director of the Environmental Economics Program at Harvard University, and University Fellow, Resources for the Future. This article draws, in part, on Revesz and Stavins (2005). Helpful comments on a previous version were provided by Lori Bennear and Nathaniel Keohane.
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 (assuming convexity of the respective functions).

The Kaldor-Hicks criterion is clearly more practical than the strict Pareto criterion, but its normative standing is less solid. 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 (Sagoff, 1993). Many economists would agree with this assertion, and some have noted that the Kaldor-Hicks criterion should be considered neither a necessary nor sufficient condition for public policy (Arrow et al., 1996).

**Benefit-Cost Analysis of Environmental Regulations**

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. The present value of net benefits (PVNB) is defined as:

$$PVNB = \sum_{t=0}^{T} \{ (B_t - C_t) \cdot (1 + r)^{-t} \}$$

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 (Goulder and Stavins, 2002).

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. In theory, the social discount rate could be derived by aggregating the individual time preference rates of all parties affected by a policy. 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 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 are subject to time inconsistency.

**The Costs of Environmental Regulations**

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. A full taxonomy of environmental costs ranges from the most obvious to the least direct (Jaffe et al., 1995).
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, such as through the use of computable general equilibrium models (Hazilla and Kopp, 1990; Conrad, 2002). The potential interaction of abatement costs with pre-existing taxes indicates the importance of employing general equilibrium models for comprehensive cost analysis. Revenue recycling (using emission tax or auctioned permit revenues to reduce distortionary taxes) can significantly lower the costs of pollution control, relative to what costs would otherwise be (Goulder, 1995).

In a retrospective examination of 28 environmental and occupational safety regulations, Harrington, Morgenstern, and Nelson (2000) found that fourteen cost estimation analyses had produced *ex ante* cost estimates that exceeded actual *ex post* costs, apparently due to technological innovation stimulated by market-based instruments (see below).

### The Benefits of Environmental Regulations

Protecting the environment usually involves active employment of capital, labor, and other scarce resources. 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 (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). Empirical evidence suggests larger than expected differences between willingness to pay and willingness to accept (Fisher, McClelland and Schulze, 1988). Theoretical explanations include psychological aversion to loss and poor substitutes for environmental amenities (Hanemann, 1991).

The benefits people derive from environmental protection can be categorized as being related to human health (mortality and morbidity), ecological impacts (both market and non-market), or materials damage. 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 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).
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, because most environmental regulations result in very small changes in individuals’ mortality risks. Hedonic wage studies, averted behavior, and contingent valuation (all discussed below) can provide estimates of marginal willingness to pay or willingness to accept for small changes in mortality risk, and such estimates can be normalized as the “value of a statistical life” (VSL).

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:

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

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 (3) would yield:

\[
VSL = \frac{12}{0.00002} = 6,000,000
\]

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. It is not the economic value of an individual life, because the VSL calculation 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.

Revealed Preference Methods of Environmental Benefit Estimation

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 (Bockstael and McConnell, 1983). 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. An additional challenge is posed by the necessity of disentangling attributes of the market good or service.

Recreational activities represent a potentially large class of benefits that are 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. 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 (Bockstael 1996). 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. *Random utility models* explicitly model the consumer’s decision to choose a particular site from among recreation locations, assessing the probability of visiting each location. Such models can be used to value changes in environmental quality by comparing decisions to visit alternative sites (Phaneuf and Smith 2004).

All recreation demand models share limitations. First, the valuation of costs depends on 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. In addition, random utility models rely on people’s perceptions of environmental quality changes. Finally, like all revealed-preference approaches, recreation demand models can be used to estimate use value only; non-use value cannot be examined.

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 (Bennear, Stavins, and Wagner, 2003). Using panel data on fishing license sales and prices, 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.

*Hedonic pricing methods* are founded on the proposition that people value goods in terms of the bundles of attributes that constitute those goods. *Hedonic property value methods* employ data on residential property values and home characteristics, including structural, neighborhood, and environmental quality attributes (Palmquist, 2003). By regressing the property value on key attributes, the hedonic price function is estimated:

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

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 (5), 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:

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

This marginal implicit price, $P_e$, measures the aggregate marginal willingness to pay for the attribute in question. For purposes of benefit estimation, the demand function for the attribute is required, and so it is necessary to examine how the marginal implicit price of the environmental
attribute varies with changes in the quantity of the attribute and other relevant variables. If the hedonic price equation (5) is non-linear, then fitted values of $P_e$ can be calculated as $e$ is varied, and a second-stage equation can be estimated:

$$\hat{P}_e = g(e, y)$$  \hspace{1cm} (7)

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

Equation (7), 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 problems. Most important among these is the question of whether a demand function has actually been estimated, since environmental quality may affect both the demand for housing and its supply, raising the classic identification problem. In addition, informational asymmetries may distort the analysis. Also, 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.

A related benefit-estimation technique is the *hedonic wage method*, based on the reality that individuals in well functioning labor markets make trade-offs between wages and risk of on-the-job injuries (or death). A job is a bundle of characteristics, including its wage, responsibilities, and risk, among others factors. 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 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:

$$W = h(\tilde{\chi}, r)$$  \hspace{1cm} (8)

where $W$ = wage (in annual terms);
$\tilde{\chi}$ = 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:

$$\frac{\partial W}{\partial r} = \frac{\partial h(\cdot)}{\partial r} = W_r$$  \hspace{1cm} (9)

This marginal implicit price of risk is the average annual income necessary to compensate a worker for a marginal change in risk throughout the year, and it 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. 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, and imperfections in labor markets (less than perfect mobility) can cause further problems.

Direct application of the method in the environmental realm 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). In such applications, the hedonic wage method brings with it possible bias, because studies typically focus on risky occupations, which may attract workers who are systematically less risk-averse.

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; Mrozek and Taylor, 2002; Viscusi and Aldy, 2003; Alberini et al., 2004).

**Stated Preference Methods of Environmental Benefit Estimation**

In the best known stated preference method, contingent valuation (CV), survey respondents are presented with scenarios that require them to trade-off, hypothetically, something for a change in an environmental good or service (Mitchell and Carson, 1989; Boyle, 2003). 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. 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 starting-point bias. A related approach is the use of a payment card shown to the respondent, but the range of WTP on the card may introduce bias, and the approach cannot be used with telephone surveys. Finally, the referendum (discrete choice) approach is favored by researchers. Each respondent is offered a different WTP number, to which a simple yes-no response is solicited.

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. 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 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).
The 1989 Exxon Valdez oil spill off the coast of Alaska led to massive litigation, and resulted in the most prominent use ever of the concept of non-use value and the method of contingent valuation for its estimation. The result was a symposium sponsored by the Exxon Corporation attacking the CV method (Hausman, 1993), and the subsequent 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. 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.

It is important to distinguish between legitimate methods of benefit estimation and approaches sometimes encountered in the policy process that do not measure willingness-to-pay or willingness-to-accept. Frequently misused techniques include: (1) employing as proxies for the benefits of a policy estimates of the “cost avoided” by not using the next most costly means of achieving the policy’s goals; (2) “societal revealed preference” models, which seek to infer the benefits of a proposed policy from the costs of previous regulatory actions; and (3) cost-of-illness or human-capital measures which estimate explicit market costs resulting from changes in morbidity or mortality. Because none of these approaches provide estimates of WTP or WTA, these techniques do not provide valid measures of economic benefits.

Choosing Instruments: The Means of Environmental Policy

Even if the goals of environmental policies are given, economic analysis can bring insights to the assessment and design of environmental policies. One important criterion is cost-effectiveness, defined as the 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:

\[
\min_{\{r_i\}} C = \sum_{i=1}^{N} c_i(r_i) \quad (10)
\]

\[
s.t. \quad \sum_{i=1}^{N} [u_i - r_i] \leq \bar{E} \quad (11)
\]

\[
\text{and} \quad 0 \leq r_i \leq u_i \quad (12)
\]

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  
\( \bar{E} = \) 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):

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

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 environmental policy instruments, a key question is whether marginal abatement costs are likely to be being equated across sources.

**Command-and-Control versus Market-Based Instruments**

Conventional approaches to regulating the environment — frequently characterized as command-and-control — allow relatively little flexibility in the means of achieving goals. Such policy instruments tend to force firms to take on equal shares of the pollution-control burden, regardless of the cost. The most prevalent form of uniform command-and-control standards are technology standards that specify the adoption of specific pollution-control technologies, and performance standards which specify uniform limits on the amount of pollution a facility can generate. In theory, non-uniform performance standards could be made to be cost-effective, but the government typically lacks the requisite information (on marginal costs of individual sources).

Market-based instruments encourage behavior through market signals, rather than through explicit directives regarding pollution control levels or methods. Market-based instruments fall within four categories: pollution charges, tradeable permits, market-friction reductions, and government subsidy reductions. Liability rules may 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 decision.

Where there is significant heterogeneity of abatement costs, command-and-control methods will not be cost-effective. 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 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).

In theory, market-based instruments allow any desired level of pollution cleanup to be realized at the lowest overall cost, 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). In addition, market-based instruments have the potential to bring down abatement costs over time 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; Maleug, 1989; Milliman and Prince, 1989; Jaffe and Stavins, 1995). However, the ranking among policy instruments, in terms of their respective impacts on technology innovation and diffusion is ambiguous (Jaffe, Newell, and Stavins, 2003).

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. Vintage-differentiated regulation is a common feature of many environmental policies, wherein the standard for regulated units is fixed in terms of their date of entry, with later vintages facing more stringent regulation. Such vintage-differentiated regulations can be expected to retard turnover in the capital stock, and thereby to reduce the cost-effectiveness of regulation. 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 (Maloney and Brady, 1988); both types of consequences have been validated empirically (Gruenspecht, 1982; Nelson, Tietenberg, Donihue, 1993).

**Pollution Charges**

Pollution charge systems assess a fee or tax on the amount of pollution that firms or sources generate (Pigou, 1920). 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:

\[
\min_{\{e_i\}} \left[ c_i(r_i) + t \cdot (u_i - r_i) \right] 
\]

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

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 are equated across sources, satisfying the condition for cost-effectiveness specified by equations (13) and (14), at least in the simplest case of a uniformly-mixed pollutant. In the non-uniformly-mixed pollutant case, where “hot spots” can be an issue, the respective cost-effective instrument is an “ambient charge.”

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 (Pigou, 1920), 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 is that 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, because firms pay both their abatement costs plus taxes on their residual emissions.

If charges are broadly defined, many applications can be identified (Stavins, 2003). Coming closest to true Pigovian are 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. Another important set of charge systems 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 disposal. A number of countries and states have implemented this approach 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. There has also been considerable use of environmental user charges, through which specific environmentally related services are funded. Examples include insurance premium taxes (Barthold, 1994). Another set of environmental charges are sales taxes on motor fuels, ozone-depleting chemicals, agricultural inputs, and low-mileage motor vehicles. Finally, tax differentiation has been used to encourage the use of renewable energy sources.

Tradeable Permit Systems

Tradeable permits can achieve the same cost-minimizing allocation as a charge system, while avoiding the problems of uncertain firm responses and the distributional consequences of taxes. Under a tradable permit system, an allowed overall level of pollution, \( E \), is established, and allocated among sources in the form of permits. Firms that keep emission levels below allotted levels may sell 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:

\[
\sum_{i=1}^{N} q_{0i} = E \tag{19}
\]

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

\[
\min_{[r_i]} [c_i(r_i) + p \cdot (u_i - r_i - q_{0i})] \tag{20}
\]
\[ s.t. \quad r_i \geq 0 \quad (21) \]

The result for each source is:

\[ \frac{\partial c_i(r_i)}{\partial r_i} - p \geq 0 \quad (22) \]

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

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. The unique cost-effective equilibrium is achieved independent of the initial allocation of permits (Montgomery, 1972), which is of great political significance.

The performance of a tradeable permit system can be adversely affected by: 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 (Montero, 2003).

Tradeable permits have been the most frequently used market-based system (U.S. Environmental Protection Agency, 2000). Significant applications include: the 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 (SO\(_2\)) allowance trading system for acid rain control (Schmalensee et al., 1998; Stavins, 1998; Carlson et al., 2000; Ellerman et al., 2000); the RECLAIM program in the Los Angeles metropolitan region (Harrison, 1999); and tradeable development rights for land use.

**Market Friction Reduction**

*Market friction reduction* can serve as a policy instrument for environmental protection. *Market creation* establishes markets for inputs or outputs associated with environmental quality. Examples of *market creation* include measures that facilitate the voluntary exchange of water rights and thus promote more efficient allocation and use of scarce water supplies (Howe, 1997), and policies that facilitate the restructuring of electricity generation and transmission. 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. These programs have been of two types. *Product labeling requirements* have been implemented to improve information sets available to consumers, while other programs have involved *reporting requirements* (Hamilton, 1995; Konar and Cohen 1997; Khanna, Quimio, and Bojilova 1998).
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 relatives 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 cutting inefficient subsidies that promote the use of fossil fuels.

Implications of Uncertainty for Instrument Choice

The dual task facing policy makers 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 widely 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.

In the environmental realm, benefit uncertainty and cost uncertainty are usually both present, with benefit uncertainty of 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, 2003). 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, 2002; Pizer, 2002), but such systems have not been adopted.

Conclusion

The growing use of economic analysis to inform environmental decision making marks greater acceptance of the usefulness of these tools to improve regulation. But debates about the normative standing of the Kaldor-Hicks criterion and the challenges inherent in making benefit-cost analysis operational will 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, there is only limited political support for broader use of benefit-cost analysis to assess proposed or existing environmental regulations. These analytical methods remain on the periphery
of policy formulation. In a growing literature (not reviewed in this article), economists have examined the processes through which political decisions regarding environmental regulation are made (Stavins 2004).

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 — may provide a model for progress with regards to analysis of the ends — the targets and goals — of public policies in this domain. The change in the former realm 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 nearly every environmental problem that is tackled, ranging from endangered species preservation to regional smog to global climate change. Of course, no individual policy instrument — whether market-based or conventional — is appropriate for all environmental problems. Which instrument 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 regulated.

Bibliography


