2 The Economic Approach: Property Rights, Externalities, and Environmental Problems

The charming landscape which I saw this morning, is indubitably made up of some twenty or thirty farms. Miller owns this field, Locke that, and Manning the woodland beyond. But none of them owns the landscape. There is a property in the horizon which no man has but he whose eye can integrate all the parts, that is, the poet. This is the best part of these men's farms, yet to this their land deeds give them no title.

—Ralph Waldo Emerson, *Nature* (1836)

Introduction

Before examining specific environmental problems and the policy responses to them, it is important that we develop and clarify the economic approach, so that we have some sense of the forest before examining each of the trees. By having a feel for the conceptual framework, it becomes easier not only to deal with individual cases but also, perhaps more importantly, to see how they fit into a comprehensive approach.

In this chapter, we develop the general conceptual framework used in economics to approach environmental problems. We begin by examining the relationship between human actions, as manifested through the economic system, and the environmental consequences of those actions. We can then establish criteria for judging the desirability of the outcomes of this relationship. These criteria provide a basis for identifying the nature and severity of environmental problems, and a foundation for designing effective policies to deal with them.

Throughout this chapter, the economic point of view is contrasted with alternative points of view. These contrasts bring the economic approach into sharper focus and stimulate deeper and more critical thinking about all possible approaches.

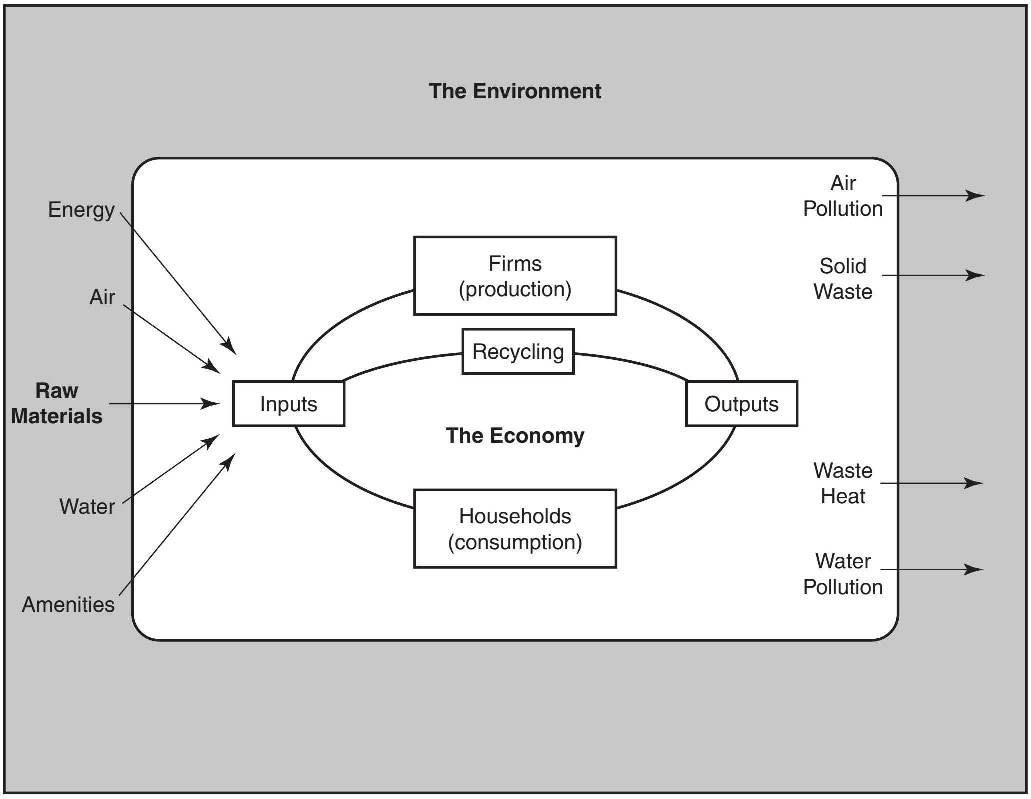
The Human–Environment Relationship

The Environment as an Asset

In economics, the environment is viewed as a composite asset that provides a variety of services. It is a very special asset, to be sure, because it provides the life-support systems that sustain our very existence, but it is an asset nonetheless. As with other assets, we wish to enhance, or at least prevent undue depreciation of, the value of this asset so that it may continue to provide aesthetic and life-sustaining services.

The environment provides the economy with raw materials, which are transformed into consumer products by the production process, and energy, which fuels this transformation. Ultimately, these raw materials and energy return to the environment as waste products (see Figure 2.1).

FIGURE 2.1 The Economic System and the Environment



The environment also provides services directly to consumers. The air we breathe, the nourishment we receive from food and drink, and the protection we derive from shelter and clothing are all benefits we receive, either directly or indirectly, from the environment. In addition, anyone who has experienced the exhilaration of white-water canoeing, the total serenity of a wilderness trek, or the breathtaking beauty of a sunset will readily recognize that the environment provides us with a variety of amenities for which no substitute exists.

If the environment is defined broadly enough, the relationship between the environment and the economic system can be considered a *closed system*. For our purposes, a closed system is one in which no inputs (energy or matter) are received from outside the system and no outputs are transferred outside the system. An *open system*, by contrast, is one in which the system imports or exports matter or energy.

If we restrict our conception of the relationship in Figure 2.1 to our planet and the atmosphere around it, then clearly we do not have a closed system. We derive most of our energy from the sun, either directly or indirectly. We have also sent spaceships well beyond the boundaries of our atmosphere. Nonetheless, historically speaking, for *material* inputs and outputs (not including energy), this system can be treated as a closed system because the amount of exports (such as abandoned space vehicles) and imports (e.g., moon rocks) are negligible. Whether the system remains closed depends on the degree to which space exploration opens up the rest of our solar system as a source of raw materials.

The treatment of our planet and its immediate environs as a closed system has an important implication that is summed up in the *first law of thermodynamics*—energy and matter can neither be created nor destroyed.[1] The law implies that the mass of materials flowing into the economic system from the environment has either to accumulate in the economic system or return to the environment as waste. When accumulation stops, the mass of materials flowing into the economic system is equal in magnitude to the mass of waste flowing into the environment.

Excessive wastes can, of course, depreciate the asset; when they exceed the absorptive capacity of nature, wastes reduce the services that the asset provides. Examples are easy to find: air pollution can cause respiratory problems; polluted drinking water can cause cancer; smog obliterates scenic vistas; climate change can lead to flooding of coastal areas.

The relationship of people to the environment is also conditioned by another physical law, the *second law of thermodynamics*. Known popularly as the *entropy law*, this law states that "entropy increases." *Entropy* is the amount of energy unavailable for work. Applied to energy processes, this law implies that no conversion from one form of energy to another is completely efficient and that the consumption of

[1]We know, however, from Einstein's famous equation ($E = mc^2$) that matter can be transformed into energy. This transformation is the source of energy in nuclear power.

energy is an irreversible process. Some energy is always lost during conversion, and the rest, once used, is no longer available for further work. The second law also implies that in the absence of new energy inputs, any closed system must eventually use up its available energy. Since energy is necessary for life, life ceases when useful energy flows cease.

We should remember that our planet is not even approximately a closed system with respect to energy; we gain energy from the sun. The entropy law does remind us, however, that the flow of solar energy establishes an upper limit on the flow of available energy that can be sustained. Once the stocks of stored energy (such as fossil fuels and nuclear energy) are gone, the amount of energy available for useful work will be determined solely by the solar flow and by the amount that can be stored (through dams, trees, and so on). Thus, in the very long run, the growth process will be limited by the availability of solar energy and our ability to put it to work.

The Economic Approach

Two different types of economic analysis can be applied to increase our understanding of the relationship between the economic system and the environment: *Positive* economics attempts to describe *what is*, *what was*, or *what will be*. *Normative* economics, by contrast, deals with what *ought to be*. Disagreements within positive economics can usually be resolved by an appeal to the facts. Normative disagreements, however, involve value judgments.

Both branches are useful. Suppose, for example, we want to investigate the relationship between trade and the environment. Positive economics could be used to describe the kinds of impacts trade would have on the economy and the environment. It could not, however, provide any guidance on the question of whether trade was desirable. That judgment would have to come from normative economics, a topic we explore in the next section.

The fact that positive analysis does not, by itself, determine the desirability of some policy action does not mean that it is not useful in the policy process. Example 2.1 provides one example of the kinds of economic impact analyses that are used in the policy process.

A rather different context for normative economics can arise when the possibilities are more open-ended. For example, we might ask, how much should we control emissions of greenhouse gases (which contribute to climate change) and how should we achieve that degree of control? Or we might ask, how much forest of various types should be preserved? Answering these questions requires us to consider the entire range of possible outcomes and to select the best or optimal one. Although that is a much more difficult question to answer than one that asks us only to compare two predefined alternatives, the basic normative analysis framework is the same in both cases.

EXAMPLE 2.1

Economic Impacts of Reducing Hazardous Pollutant Emissions from Iron and Steel Foundries

The U.S. Environmental Protection Agency (EPA) was tasked with developing a "maximum achievable control technology standard" to reduce emissions of hazardous air pollutants from iron and steel foundries. As part of the rule-making process, EPA conducted an *ex ante* economic impact analysis to assess the potential economic impacts of the proposed rule.

If implemented, the rule would require some iron and steel foundries to implement pollution control methods that would increase the production costs at affected facilities. The interesting question addressed by the analysis is how large those impacts would be.

The impact analysis estimated annual costs for existing sources to be $21.73 million. These cost increases were projected to result in small increases in output prices. Specifically, prices were projected to increase by only 0.1 percent for iron castings and 0.05 percent for steel castings. The impacts of these price increases were expected to be experienced largely by iron foundries using cupola furnaces as well as consumers of iron foundry products. Unaffected domestic foundries and foreign producers of coke were actually projected to earn slightly higher profits as a result of the rule.

This analysis helped in two ways. First, by showing that the impacts fell under the $100 million threshold that mandates review by the Office of Management and Budget, the analysis eliminated the need for a much more time and resource consuming analysis. Second, by showing how small the expected impacts would be, it served to lower the opposition that might have arisen from unfounded fears of much more severe impacts.

Source: Office of Air Quality Planning and Standards, United States Environmental Protection Agency, "Economic Impact Analysis of Proposed Iron and Steel Foundries." NESHAP Final Report, November 2002; and National Emissions Standards for Hazardous Air Pollutants for Iron and Steel Foundries, Proposed Rule, FEDERAL REGISTER, Vol. 72, No. 73 (April 17, 2007), pp 19150–19164.

Environmental Problems and Economic Efficiency

Static Efficiency

The chief normative economic criterion for choosing among various outcomes occurring at the same point in time is called *static efficiency*, or merely *efficiency*. An allocation of resources is said to satisfy the static efficiency criterion if the economic surplus derived from those resources is maximized by that allocation. Economic surplus, in turn, is the sum of consumer's surplus and producer's surplus.

Consumer surplus is the value that consumers receive from an allocation minus what it costs them to obtain it. Consumer surplus is measured as the area under the demand curve minus the consumer's cost. The cost to the consumer is the area under the price line, bounded from the left by the vertical axis and the right by the quantity of the good. This rectangle, which captures price times quantity, represents consumer expenditure on this quantity of the good.

Why is this area thought of as a surplus? For each quantity purchased, the corresponding point on the market demand curve represents the amount of money some person would have been willing to pay for the last unit of the good. The *total willingness to pay* for some quantity of this good—say, three units—is the sum of the willingness to pay for each of the three units. Thus, the total willingness to pay for three units would be measured by the sum of the willingness to pay for the first, second, and third units, respectively. It is now a simple extension to note that the total willingness to pay is the area under the continuous market demand curve to the left of the allocation in question. For example, in Figure 2.2 the total willingness to pay for Q_d units of the commodity is the shaded area. Total willingness to pay is the concept we shall use to define the total value a consumer would receive from the five units of the good. Thus, total value the consumer would receive is equal to the area under the market demand curve from the origin to the allocation of interest. Consumer surplus is thus the excess of total willingness to pay over the (lower) actual cost.

Meanwhile, sellers face a similar choice (see Figure 2.3). Given price P^*, the seller maximizes his or her own producer surplus by choosing to sell Q_s units. The

FIGURE 2.2 The Consumer's Choice

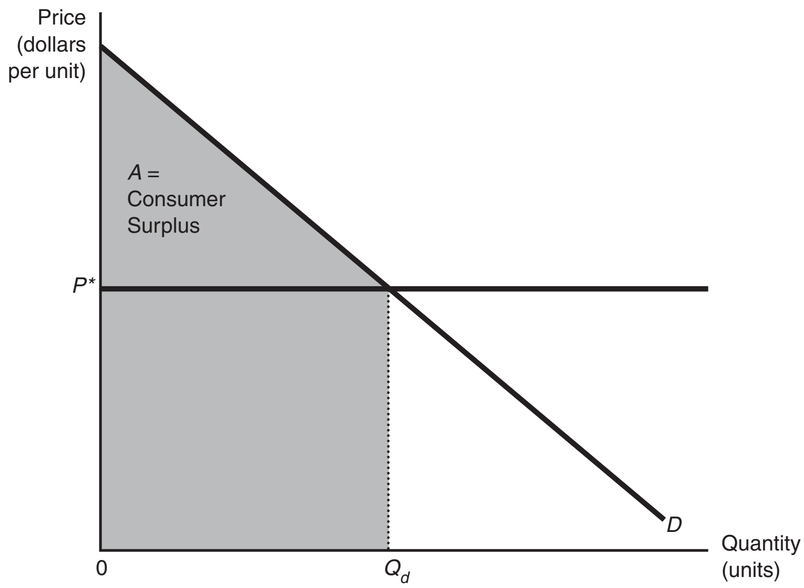
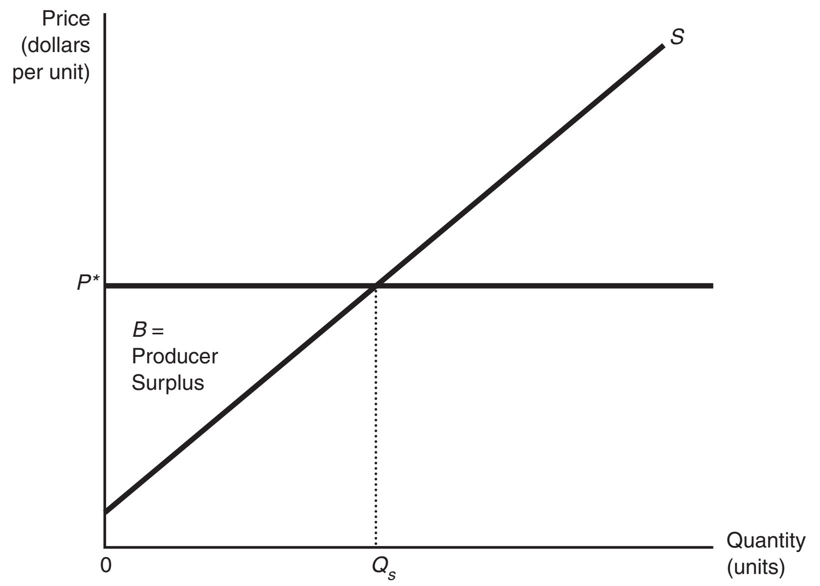


FIGURE 2.3 The Producer's Choice



producer surplus is designated by area *B*, the area under the price line that lies over the marginal cost curve, bounded from the left by the vertical axis and the right by the quantity of the good.

Property Rights

Property Rights and Efficient Market Allocations

The manner in which producers and consumers use environmental resources depends on the property rights governing those resources. In economics, *property right* refers to a bundle of entitlements defining the owner's rights, privileges, and limitations for use of the resource. By examining such entitlements and how they affect human behavior, we will better understand how environmental problems arise from government and market allocations.

These property rights can be vested either with individuals, as in a capitalist economy, or with the state, as in a centrally planned socialist economy. How can we tell when the pursuit of profits is consistent with efficiency and when it is not?

Efficient Property Rights Structures

Let's begin by describing the structure of property rights that could produce efficient allocations in a well-functioning market economy. An efficient structure has three main characteristics:

1. *Exclusivity*—All benefits and costs accrued as a result of owning and using the resources should accrue to the owner, and only to the owner, either directly or indirectly by sale to others.
2. *Transferability*—All property rights should be transferable from one owner to another in a voluntary exchange.
3. *Enforceability*—Property rights should be secure from involuntary seizure or encroachment by others.

An owner of a resource with a well-defined property right (one exhibiting these three characteristics) has a powerful incentive to use that resource efficiently because a decline in the value of that resource represents a personal loss. Farmers who own the land have an incentive to fertilize and irrigate it because the resulting increased production raises income. Similarly, they have an incentive to rotate crops when that raises the productivity of their land.

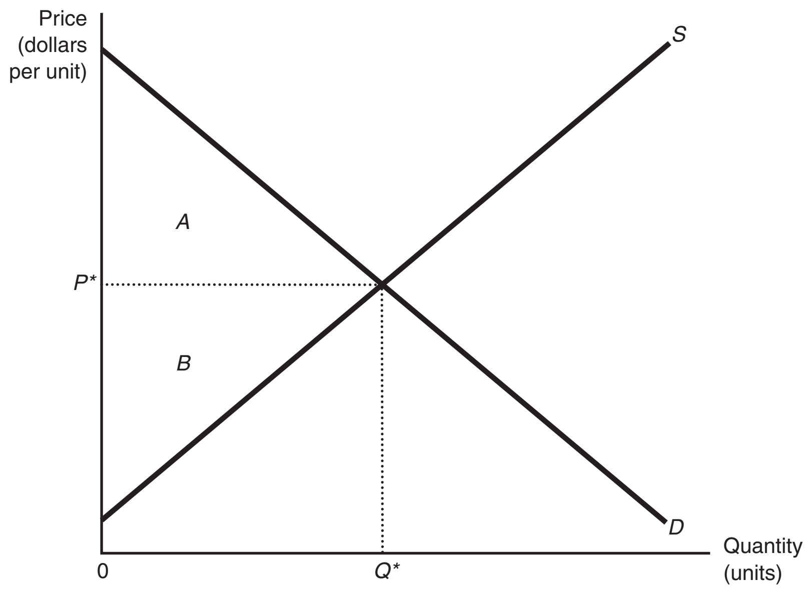
When well-defined property rights are exchanged, as in a market economy, this exchange facilitates efficiency. We can illustrate this point by examining the incentives consumers and producers face when a well-defined system of property rights is in place. Because the seller has the right to prevent the consumer from consuming the product in the absence of payment, the consumer must pay to receive the product. Given a market price, the consumer decides how much to purchase by choosing the amount that maximizes his or her individual consumer surplus.

Is this allocation efficient? According to our definition of static efficiency, it is clear the answer is yes. The economic surplus is maximized by the market allocation and, as seen in Figure 2.4, it is equal to the sum of consumer and producer surpluses (areas $A + B$). Thus, we have established a procedure for measuring efficiency, and a means of describing how the surplus is distributed between consumers and producers.

This distinction is crucially significant. Efficiency is *not* achieved because consumers and producers are seeking efficiency. They aren't! In a system with well-defined property rights and competitive markets in which to sell those rights, producers try to maximize their surplus and consumers try to maximize their surplus. The price system, then, induces those self-interested parties to make choices that are efficient from the point of view of society as a whole. It channels the energy motivated by self-interest into socially productive paths.

Familiarity may have dulled our appreciation, but it is noteworthy that a system designed to produce a harmonious and congenial outcome could function effectively while allowing consumers and producers so much individual freedom in making choices. This is truly a remarkable accomplishment.

FIGURE 2.4 Market Equilibrium



Producer's Surplus, Scarcity Rent, and Long-Run Competitive Equilibrium

Since the area under the price line is total revenue, and the area under the marginal cost curve is total variable cost, producer's surplus is related to profits. In the short run when some costs are fixed, producer's surplus is equal to profits plus fixed cost. In the long run when all costs are variable, producer's surplus is equal to profits plus rent, the return to scarce inputs owned by the producer. As long as new firms can enter into profitable industries without raising the prices of purchased inputs, long-run profits and rent will equal zero.

Scarcity Rent. Most natural resource industries, however, do give rise to rent and, therefore, producer's surplus is not eliminated by competition, even with free entry. This producer's surplus, which persists in long-run competitive equilibrium, is called *scarcity rent.*

David Ricardo was the first economist to recognize the existence of scarcity rent. Ricardo suggested that the price of land was determined by the least fertile marginal unit of land. Since the price had to be sufficiently high to allow the poorer land to be brought into production, other, more fertile land could be farmed at an economic profit. Competition could not erode that profit because the amount of high quality land was limited and lower prices would serve only to reduce the supply of land below demand. The only way to expand production would be to bring additional, less fertile land (more costly to farm) into production; consequently, additional production does not lower price, as it does in a constant-cost industry. As we shall see, other circumstances also give rise to scarcity rent for natural resources.

Externalities as a Source of Market Failure

The Concept Introduced

Exclusivity is one of the chief characteristics of an efficient property rights structure. This characteristic is frequently violated in practice. One broad class of violations occurs when an agent making a decision does not bear all of the consequences of his or her action.

Suppose two firms are located by a river. The first produces steel, while the second, somewhat downstream, operates a resort hotel. Both use the river, although in different ways. The steel firm uses it as a receptacle for its waste, while the hotel uses it to attract customers seeking water recreation. If these two facilities have different owners, an efficient use of the water is not likely to result. Because the steel plant does not bear the cost of reduced business at the resort resulting from waste being dumped into the river, it is not likely to be very sensitive to that cost in its decision making. As a result, it could be expected to dump too much waste into the river, and an efficient allocation of the river would not be attained.

This situation is called an externality. An *externality* exists whenever the welfare of some agent, either a firm or household, depends not only on his or her activities, but also on activities under the control of some other agent. In the example, the increased waste in the river imposed an external cost on the resort, a cost the steel firm could not be counted upon to consider appropriately in deciding the amount of waste to dump.

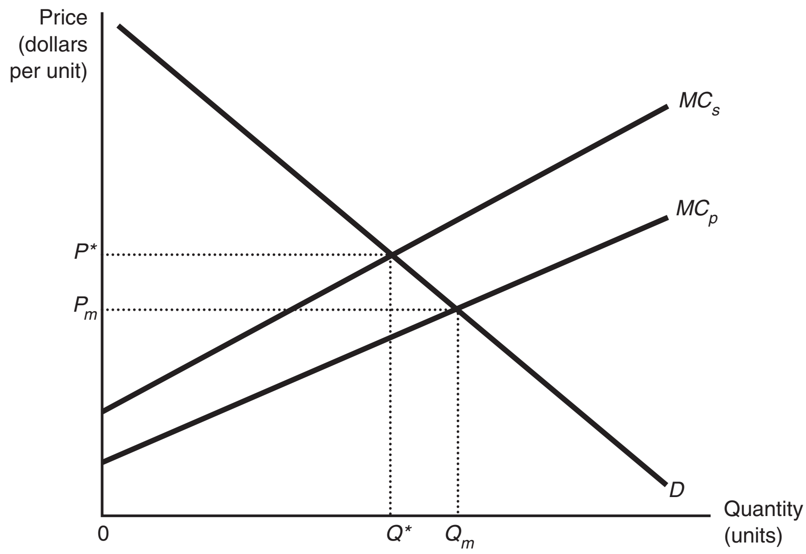
The effect of this external cost on the steel industry is illustrated in Figure 2.5, which shows the market for steel. Steel production inevitably involves producing pollution as well as steel. The demand for steel is shown by the demand curve D, and the private marginal cost of producing the steel (exclusive of pollution control and damage) is depicted as MC_p. Because society considers both the cost of pollution and the cost of producing the steel, the social marginal cost function (MC_s) includes both of these costs as well.

If the steel industry faced no outside control on its emission levels, it would seek to produce Q_m. That choice, in a competitive setting, would maximize its private producer surplus. But that is clearly not efficient, since the net benefit is maximized at Q^*, not Q_m.

With the help of Figure 2.5, we can draw a number of conclusions about market allocations of commodities causing pollution externalities:

1. The output of the commodity is too large.
2. Too much pollution is produced.

FIGURE 2.5 The Market for Steel



3. The prices of products responsible for pollution are too low.
4. As long as the costs are external, no incentives to search for ways to yield less pollution per unit of output are introduced by the market.
5. Recycling and reuse of the polluting substances are discouraged because release into the environment is so inefficiently cheap.

The effects of a market imperfection for one commodity end up affecting the demands for raw materials, labor, and so on. The ultimate effects are felt through the entire economy.

Types of Externalities

External effects, or externalities, can be positive or negative. Historically, the terms *external diseconomy* and *external economy* have been used to refer, respectively, to circumstances in which the affected party is damaged by or benefits from the externality. Clearly, the water pollution example represents an external diseconomy. External economies are not hard to find, however. Private individuals who preserve a particularly scenic area provide an external economy to all who pass. Generally, when external economies are present, the market will undersupply the resources.

One other distinction is important. One class of externalities, known as *pecuniary externalities*, does not present the same kinds of problems as pollution does. Pecuniary externalities arise when the external effect is transmitted through altered prices. Suppose that a new firm moves into an area and drives up the rental price of land. That increase creates a negative effect on all those paying rent and, therefore, is an external diseconomy.

This pecuniary diseconomy, however, does not cause a market failure because the resulting higher rents are reflecting the scarcity of land. The land market provides a mechanism by which the parties can bid for land; the resulting prices reflect the value of the land in its various uses. Without pecuniary externalities, the price signals would fail to sustain an efficient allocation.

The pollution example is *not* a pecuniary externality because the effect is not transmitted through prices. In this example, prices do not adjust to reflect the increasing waste load. The damage to the water resource is not reflected in the steel firm's costs. An essential feedback mechanism that is present for pecuniary externalities is not present for the pollution case.

The externalities concept is a broad one, covering a multitude of sources of market failure (Example 2.2 illustrates one). The next step is to investigate some specific circumstances that can give rise to externalities.

EXAMPLE 2.2

Shrimp Farming Externalities in Thailand

In the Tha Po village on the coast of Surat Thani Province in Thailand, more than half of the 1,100 hectares of mangrove swamps have been cleared for commercial shrimp farms. Although harvesting shrimp is a lucrative undertaking, mangroves serve as nurseries for fish and as barriers for storms and soil erosion. Following the destruction of the local mangroves, Tha Po villagers experienced a decline in fish catch and suffered storm damage and water pollution. Can market forces be trusted to strike the efficient balance between preservation and development for the remaining mangroves?

Calculations by economists Sathirathai and Barbier (2001) demonstrated that the value of the ecological services that would be lost from further destruction of the mangrove swamps exceeded the value of the shrimp farms that would take their place. Preservation of the remaining mangrove swamps would be the efficient choice.

Would a potential shrimp-farming entrepreneur make the efficient choice? Unfortunately, the answer is no. This study estimated the economic value of mangroves in terms of local use of forest resources, offshore fishery linkages, and coastal protection to be in the range of $27,264–$35,921 per hectare. In contrast, the economic returns to shrimp farming, once they are corrected for input subsidies and for the costs of water pollution, are only $194–$209 per hectare. However, as shrimp farmers are heavily subsidized and do not have to take into account the external costs of pollution, their financial returns are typically $7,706.95–$8,336.47 per hectare. In the absence of some sort of external control imposed by collective action, development would be the normal, if inefficient, result. The externalities associated with the ecological services provided by the mangroves support a biased decision that results in fewer social net benefits, but greater private net benefits.

Source: Suthawan Sathirathai and Edward B. Barbier. "Valuing Mangrove Conservation in Southern Thailand" CONTEMPORARY ECONOMIC POLICY, Vol. 19, No. 2 (April 2001), pp. 109–122.

Improperly Designed Property Rights Systems

Other Property Rights Regimes[2]

Private property is, of course, not the only possible way of defining entitlements to resource use. Other possibilities include state-property regimes (where the government owns and controls the property), common-property regimes (where the property is jointly owned and managed by a specified group of co-owners), and *res nullius* or open-access regimes (in which no one owns or exercises control over the resources). All of these create rather different incentives for resource use.

State-property regimes exist not only in former communist countries, but also to varying degrees in virtually all countries of the world. Parks and forests, for example, are frequently owned and managed by the government in capitalist as well as in socialist nations. Problems with both efficiency and sustainability can arise in state-property regimes when the incentives of bureaucrats, who implement and/or make the rules for resource use, diverge from collective interests.

Common-property resources are those shared resources that are managed in common rather than privately. Entitlements to use common-property resources may be formal, protected by specific legal rules, or they may be informal, protected by tradition or custom. Common-property regimes exhibit varying degrees of efficiency and sustainability, depending on the rules that emerge from collective decision making. While some very successful examples of common-property regimes exist, unsuccessful examples are even more common.[3]

One successful example of a common-property regime involves the system of allocating grazing rights in Switzerland. Although agricultural land is normally treated as private property in Switzerland, grazing rights on the Alpine meadows have been treated as common property for centuries. Overgrazing is protected by specific rules, enacted by an association of users, which limit the amount of livestock permitted on the meadow. The families included on the membership list of the association have been stable over time as rights and responsibilities have passed from generation to generation. This stability has apparently facilitated reciprocity and trust, thereby providing a foundation for continued compliance with the rules.

Unfortunately, that kind of stability may be the exception rather than the rule, particularly in the face of heavy population pressure. The more common situation can be illustrated by the experience of Mawelle, a small fishing village in Sri Lanka. Initially, a complicated but effective rotating system of fishing rights was devised by villagers to assure equitable access to the best spots and best times while protecting the fish stocks. Over time, population pressure and the infusion of outsiders raised demand and undermined the collective cohesion sufficiently that the traditional rules became unenforceable, producing overexploitation of the resource and lower incomes for all the participants.

---

[2]This section relies on the classification system presented in Bromley (1991).
[3]The two cases that follow, and many others, are discussed in Ostrom (1990).

Res nullius property resources, the main focus of this section, can be exploited on a first-come, first-served basis because no individual or group has the legal power to restrict access. *Open-access resources*, as we shall henceforth call them, have given rise to what has become known popularly as the "tragedy of the commons."

The problems created by open-access resources can be illustrated by recalling the fate of the American bison. Bison are an example of "common-pool" resources. *Common-pool resources* are shared resources characterized by nonexclusivity and divisibility. Nonexclusivity implies that resources can be exploited by anyone, while divisibility means that the capture of part of the resource by one group subtracts it from the amount available to the other groups. (Note the contrast between common-pool resources and public goods, the subject of the next section.) In the early history of the United States, bison were plentiful; unrestricted hunting access was not a problem. Frontier people who needed hides or meat could easily get whatever they needed; the aggressiveness of any one hunter did not affect the time and effort expended by other hunters. In the absence of scarcity, efficiency was not threatened by open access.

As the years slipped by, however, the demand for bison increased and scarcity became a factor. As the number of hunters increased, eventually every additional unit of hunting activity increased the amount of time and effort required to produce a given yield of bison.

Consider graphically how various property rights structures (and the resulting level of harvest) affect the scarcity rent (in this case, equivalent to the economic surplus received by consumers and producers), where the amount of rent is measured as the difference between the revenues received from the harvest minus the costs associates with producing that harvest. Figure 2.6 compares the revenue and costs for various levels of harvest. In the top panel the revenue is calculated by multiplying, for each level of hunting activity, the (assumed constant) price of bison by the amount harvested. The upward sloping total cost curve simply reflects that fact that increases in harvest effort result in higher costs. (Marginal cost is assumed to be constant for this example.)

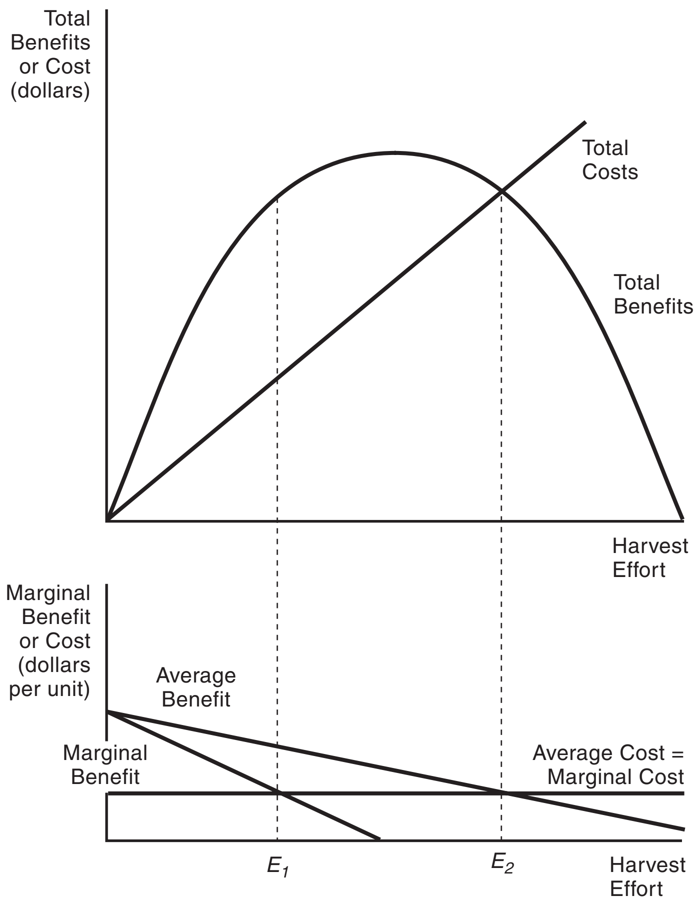
In terms of the top panel of Figure 2.6 the total surplus associated with any level of effort is measured as the vertical difference between the total revenue curve and the total cost curve for that level of harvest.

In the bottom panel the marginal revenue curve is downward sloping (despite the constant price) because as the amount of hunting effort increases, the resulting bison population size decreases. Smaller populations support smaller harvests per unit of effort expended.

The efficient level of hunting activity in this model (E_1) maximizes the surplus. This can be seen graphically in two different ways. First, E_1 maximizes the vertical difference between the two curves in the top panel. Second, in the bottom panel E_1 is the level where the marginal revenue, which records the addition to the surplus from an additional unit of effort, crosses the marginal cost curve, which measures the reduction in the surplus due to the additional cost of expending that last unit of effort. These are simply two different (mathematically equivalent) ways to demonstrate the same outcome. (The curves in the bottom panel are derived from the curves in the top panel.)

With all hunters having completely unrestricted access to the bison, the resulting allocation would not be efficient. No individual hunter would have an incentive

FIGURE 2.6 Bison Harvesting



to protect scarcity rent by restricting hunting effort. Individual hunters, without exclusive rights, would exploit the resource until their total benefit equaled total cost, implying a level of effort equal to (E_2). Excessive exploitation of the herd occurs because individual hunters cannot appropriate the scarcity rent; therefore, they ignore it. One of the losses from further exploitation that could be avoided by exclusive owners—the loss of scarcity rent due to overexploitation—is not part of the decision-making process of open-access hunters.

Two characteristics of this formulation of the open-access allocation are worth noting: (1) In the presence of sufficient demand, unrestricted access will cause resources to be overexploited; (2) the scarcity rent is dissipated; no one is able to appropriate the rent, so it is lost.

Why does this happen? Unlimited access destroys the incentive to conserve. A hunter who can preclude others from hunting his stock has an incentive to keep the herd at an efficient level. This restraint results in lower costs in the form of less time and effort expended to produce a given yield of bison. On the other hand, a hunter exploiting an open-access resource would not have an incentive to conserve because the potential additional economic surplus derived from self-restraint would, to some extent, be captured by other hunters who simply kept harvesting. Thus, unrestricted access to resources promotes an inefficient allocation. As a result of excessive harvest and the loss of habitat as land was converted to farm and pasture, the Great Plains bison herds nearly became extinct (Lueck, 2002). Another example of open-access, fisheries, is the principal topic of Chapter 13.

Public Goods

Public goods, defined as those that exhibit both consumption indivisibilities and nonexcludability, present a particularly complex category of environmental resources. *Nonexcludability* refers to a circumstance where, once the resource is provided, even those who fail to pay for it cannot be excluded from enjoying the benefits it confers. Consumption is said to be *indivisible* when one person's consumption of a good does not diminish the amount available for others. Several common environmental resources are public goods, including not only the "charming landscape" referred to by Emerson, but also clean air, clean water, and biological diversity.[4]

Biological diversity includes two related concepts: (1) the amount of genetic variability among individuals within a single species, and (2) the number of species within a community of organisms. *Genetic diversity*, critical to species survival in the natural world, has also proved to be important in the development of new crops and livestock. It enhances the opportunities for crossbreeding and, thus, the development of superior strains. The availability of different strains was the key, for example, in developing new, disease-resistant barley.

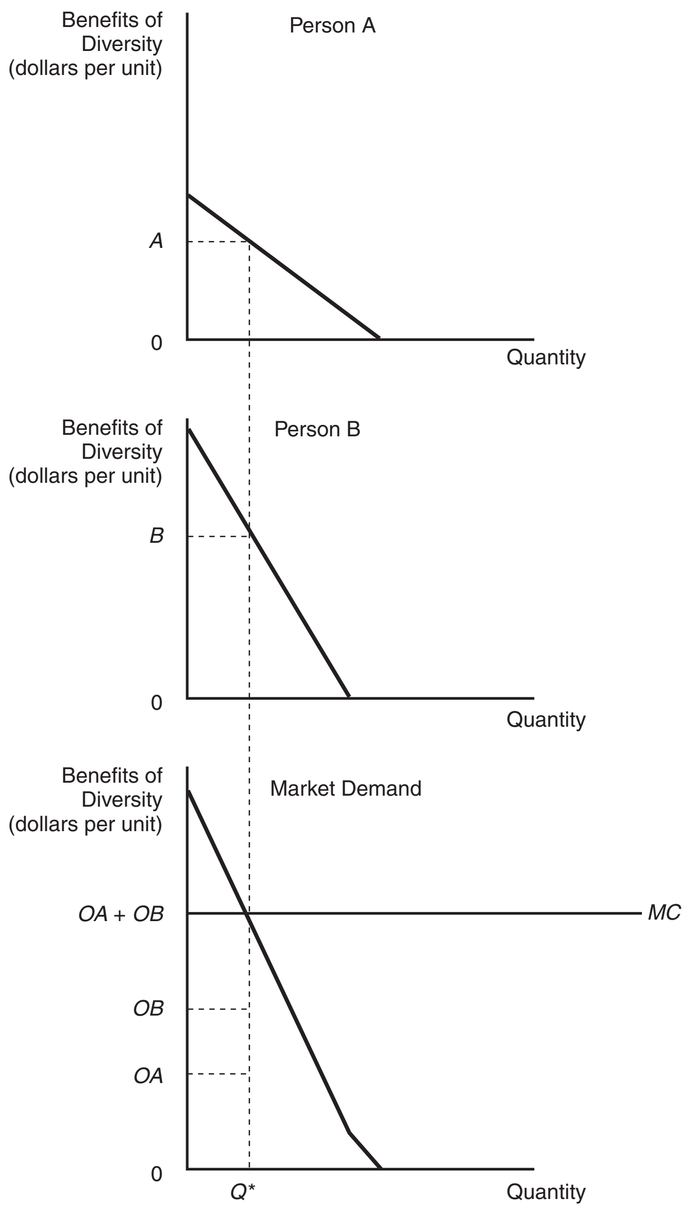
Because of the interdependence of species within ecological communities, any particular species may have a value to the community far beyond its intrinsic value. Certain species contribute balance and stability to their ecological communities by providing food sources or holding the population of the species in check.

The richness of diversity within and among species has provided new sources of food, energy, industrial chemicals, raw materials, and medicines. Yet, considerable evidence suggests that biological diversity is decreasing.

Can we rely on the private sector to produce the efficient amount of public goods, such as biological diversity? Unfortunately, the answer is no! Suppose that in response to diminishing ecological diversity we decide to take up a collection to provide some means of preserving endangered species. Would the collection yield

[4]Notice that public "bads," such as dirty air and dirty water, are also possible.

FIGURE 2.7 Efficient Provision of Public Goods



sufficient revenue to pay for an efficient level of ecological diversity? The general answer is no. Let's see why.

In Figure 2.7, individual demand curves for preserving biodiversity have been presented for two consumers A and B. The market demand curve is represented by the vertical summation of the two individual demand curves. A vertical summation is necessary because everyone can simultaneously consume the same amount of biological diversity. We are, therefore, able to determine the market demand by finding the sum of the amounts of money they would be willing to pay for that level of diversity.

What is the efficient level of diversity? It can be determined by a direct application of our definition of efficiency. The efficient allocation maximizes economic surplus, which is represented geometrically by the portion of the area under the market demand curve that lies above the constant marginal cost curve. The allocation that maximizes economic surplus is Q^*, the allocation where the demand curve crosses the marginal cost curve.

Why would a competitive market not be expected to supply the efficient level of this good? Since the two consumers get very different marginal willingness to pay from the efficient allocation of this good (*OA* versus *OB*), the efficient pricing system would require charging a different price to each consumer. Person *A* would pay *OA* and person *B* would pay *OB*. (Remember consumers tend to choose the level of the good that equates their marginal willingness to pay to the price they face.) Yet the producer would have no basis for figuring out how to differentiate the prices. In the absence of excludability, consumers are not likely choose to reveal the strength of their preference for this commodity. All consumers have an incentive to understate the strength of their preferences to try to shift more of the cost burden to the other consumers.

Therefore, inefficiency results because each person is able to become a free rider on the other's contribution. A *free rider* is someone who derives the value from a commodity without paying an efficient amount for its supply. Because of the consumption indivisibility and nonexcludability properties of the public good, consumers receive the value of any diversity purchased by other people. When this happens it tends to diminish incentives to contribute, and the contributions are not sufficiently large to finance the efficient amount of the public good; it would be undersupplied.

The privately supplied amount may not be zero, however. Some diversity would be privately supplied. Indeed, as suggested by Example 2.3, the privately supplied amount may be considerable.

Imperfect Market Structures

Environmental problems also occur when one of the participants in an exchange of property rights is able to exercise an inordinate amount of power over the outcome. This can occur, for example, when a product is sold by a single seller, or *monopoly*.

It is easy to show that monopolies violate our definition of *efficiency* in the goods market (see Figure 2.8). According to our definition of *static efficiency*, the efficient allocation would result when *OB* is supplied. This would yield consumer surplus represented by triangle *IGC and* producer surplus denoted by triangle *GCH*.

EXAMPLE 2.3

Public Goods Privately Provided: The Nature Conservancy

Can the demand for a public good such as biological diversity be observed in practice? Would the market respond to that demand? Apparently so, according to the existence of an organization called The Nature Conservancy.

The Nature Conservancy was born of an older organization called the Ecologist Union on September 11, 1950, for the purpose of establishing natural area reserves to aid in the preservation of areas, objects, and fauna and flora that have scientific, educational, or aesthetic significance. This organization purchases, or accepts as donations, land that has some unique ecological or aesthetic significance, to keep it from being used for other purposes. In so doing they preserve many species by preserving the habitat.

From humble beginnings, The Nature Conservancy has, as of 2010, been responsible for the preservation of 119 million acres of forests, marshes, prairies, mounds, and islands around the world. Additionally, The Nature Conservancy has protected 5,000 miles of rivers and operates over 100 marine conservation projects. These areas serve as home to rare and endangered species of wildlife and plants. The Conservancy owns and manages the largest privately owned nature preserve system in the world.

This approach has considerable merit. A private organization can move more rapidly than the public sector. Because it has a limited budget, The Nature Conservancy sets priorities and concentrates on acquiring the most ecologically unique areas. Yet the theory of public goods reminds us that if this were to be the sole approach to the preservation of biological diversity, it would preserve a smaller-than-efficient amount.

Source: The Nature Conservancy, http://nature.org/aboutus/.

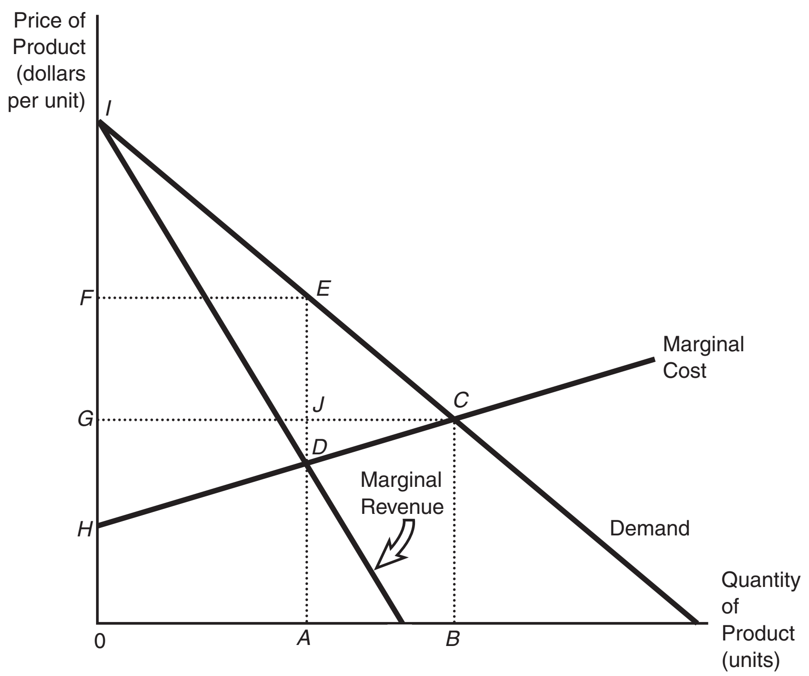
The monopoly, however, would produce and sell *OA*, where marginal revenue equals marginal cost, and would charge price *OF.* At this point, although the producer's surplus (HFED) is maximized, the sum of consumer and producer surplus is clearly not, because this choice causes society to lose economic surplus equal to triangle *EDC.*[5] Monopolies supply an inefficiently small amount of the good.

Imperfect markets clearly play some role in environmental problems. For example, the major oil-exporting countries have formed a cartel, resulting in higher-than-normal prices and lower-than-normal production. A *cartel* is a collusive agreement among producers to restrict production and raise prices. This collusive agreement allows the group to act as a monopolist. The inefficiency in the goods market would normally be offset to some degree by the reduction in

---

[5]Producers would lose area *JDC* compared to the efficient allocation, but they would gain area *FEJG*, which is much larger. Meanwhile, consumers would be worse off, because they lose area *FECJG*. Of these, *FEJG* is merely a transfer to the monopoly, whereas *EJC* is a pure loss to society. The total pure loss (*EDC*) is called a *deadweight loss*.

FIGURE 2.8 Monopoly and Inefficiency



social costs caused by the lower levels of pollution resulting from the reduction in the combustion of oil. Debate 2.1 examines the pricing activities of OPEC and recent fluctuations in oil prices.

Government Failure

Market processes are not the only sources of inefficiency. Political processes are fully as culpable. As will become clear in the chapters that follow, some environmental problems have arisen from a failure of political, rather than economic, institutions. To complete our study of the ability of institutions to allocate environmental resources, we must understand this source of inefficiency as well.

Government failure shares with market failure the characteristic that improper incentives are the root of the problem. Special interest groups use the political process to engage in what has become known as *rent seeking*. Rent seeking is the use of resources in lobbying and other activities directed at securing protective legislation. Successful rent-seeking activity will increase the net benefits going to the special interest group, but it will also frequently lower the surplus to society as a whole. In these instances, it is a classic case of the aggressive pursuit of a larger slice of the pie leading to a smaller pie.

DEBATE 2.1

How Should OPEC Price Its Oil?

As a cartel, OPEC (Organization of Petroleum Exporting Countries) has considerable control over its output and, hence, prices. And as Figure 2.8 suggests, it could increase its profits by restricting supply, a tactic that would cause prices to rise above their competitive levels. By how much should prices be raised?

The profit-maximizing price will depend upon several factors, including the price elasticity of demand (to determine how much the quantity demanded will fall in response to the higher price), the price elasticity of supply for non-OPEC members (to determine how much added production should be expected from outside producers), and the propensity for cheating (members producing more than their assigned quotas). Gately (1995) has modeled these and other factors and concluded that OPEC's interests would be best served by a policy of moderate output growth, defined as growth at a rate no faster than world income growth.

As Gately points out, however, OPEC historically has not always exercised this degree of caution. In 1979–1980, succumbing to the lure of even higher prices, OPEC chose a price strategy that required substantial restrictions of cartel output. Not only did the price elasticities of demand and non-OPEC supply turn out to be much higher than anticipated by the cartel, but also the higher oil prices triggered a worldwide recession (which further lowered demand). OPEC lost not only revenue but also market share. Even for monopolies, the market imposes some discipline; the highest price is not always the best price.

Interestingly, since 1980, world oil markets have experienced increasing price volatility. Oil prices dropped as low as $10 per barrel in 1998 and rose above $30 per barrel in 2000 (then considered a huge price swing). In 2008, oil prices rose to over $138 per barrel! Kohl (2002) analyzes OPEC's behavior during the period of 1998–2001. He notes that OPEC has consistently had trouble with member compliance and with the non-OPEC competitive fringe (e.g., Norway, Mexico, and Russia). He notes that compliance with production quotas has been best during periods of high demand or when the quotas are set above production capacity.

High demand has been the recent situation. With surging demand in China and the United States, oil prices have risen dramatically. Will higher prices induce sufficient reductions in consumption to moderate OPEC power? Stay tuned.

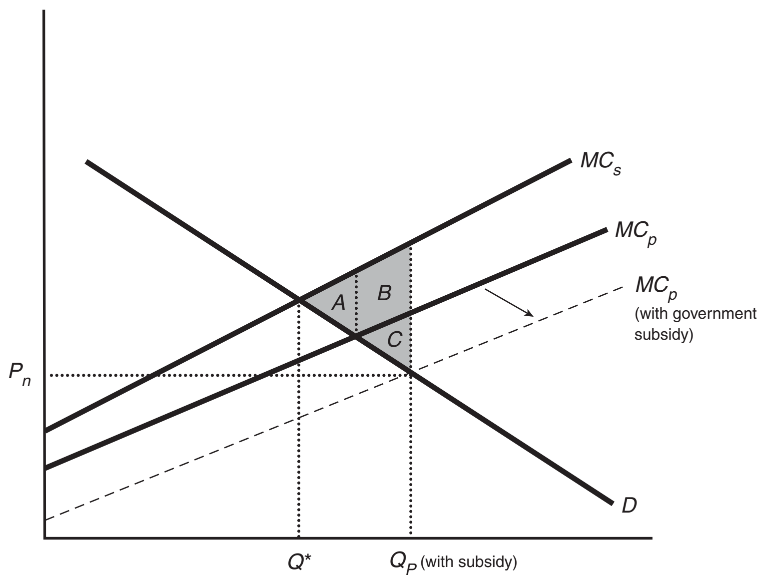
Sources: Dermot Gately. "Strategies for OPEC's Pricing and Output Decisions." ENERGY JOURNAL, Vol. 16, No. 3 (1995), pp. 1–38; Wilfrid L. Kohl, "OPEC Behavior, 1998–2001." QUARTERLY REVIEW OF ECONOMICS AND FINANCE, Vol. 42 (2002), pp. 209–233; and "OPEC Finds Price Range to Live With." THE NEW YORK TIMES, December 6, 2007.

Why don't the losers rise up to protect their interests? One main reason is voter ignorance. It is economically rational for voters to remain ignorant on many issues simply because of the high cost of keeping informed and the low probability that any single vote will be decisive. In addition, it is difficult for diffuse groups of individuals, each of whom is affected only to a small degree, to organize a coherent, unified opposition. Successful opposition is, in a sense, a public good, with its attendant tendency for free riding on the opposition of others. Opposition to special interests would normally be underfunded.

Rent seeking can take many forms. Producers can seek protection from competitive pressures brought by imports or can seek price floors to hold prices above their efficient levels. Consumer groups can seek price ceilings or special subsidies to transfer part of their costs to the general body of taxpayers. Rent seeking is not the only source of inefficient government policy. Sometimes governments act without full information and establish policies that are ultimately very inefficient. For example, as we will discuss in Chapter 17, one technological strategy chosen by the government to control motor vehicle pollution involved adding the chemical substance MTBE to gasoline. Designed to promote cleaner combustion, this additive turned out to create a substantial water pollution problem.

Governments may also pursue social policy objectives that have the side effect of causing an environmental inefficiency. For example, looking back at Figure 2.5, suppose that the government, for reasons of national security, decides to subsidize the production of steel. Figure 2.9 illustrates the outcome. The private marginal cost curve shifts down and to the right causing a further increase in production, lower prices, and even more pollution produced. Thus, the subsidy moves us even further away from where surplus is maximized at Q^*. The shaded triangle A shows the deadweight loss (inefficiency) without the subsidy. With the subsidy, the deadweight loss grows to areas $A + B + C$. This social policy has the side effect of increasing an environmental inefficiency. In another example, in Chapter 7, we shall see how the desire to hold down natural gas prices for consumers led to massive

FIGURE 2.9 The Market for Steel Revisited



shortages. These examples provide a direct challenge to the presumption that more direct intervention by the government automatically leads to either greater efficiency or greater sustainability.

These cases illustrate the general economic premise that environmental problems arise because of a divergence between individual and collective objectives. This is a powerful explanatory device because not only does it suggest why these problems arise, but also it suggests how they might be resolved—by realigning individual incentives to make them compatible with collective objectives. As self-evident as this approach may be, it is controversial. The controversy involves whether the problem is our improper values or the improper translation of our quite proper values into action.

Economists have always been reluctant to argue that values of consumers are warped, because that would necessitate dictating the "correct" set of values. Both capitalism and democracy are based on the presumption that the majority knows what it is doing, whether it is casting ballots for representatives or dollar votes for goods and services.

The Pursuit of Efficiency

We have seen that environmental problems arise when property rights are ill defined, and when these rights are exchanged under something other than competitive conditions. We can now use our definition of efficiency to explore possible remedies, such as private negotiation, judicial remedies, and regulation by the legislative and executive branches of government.

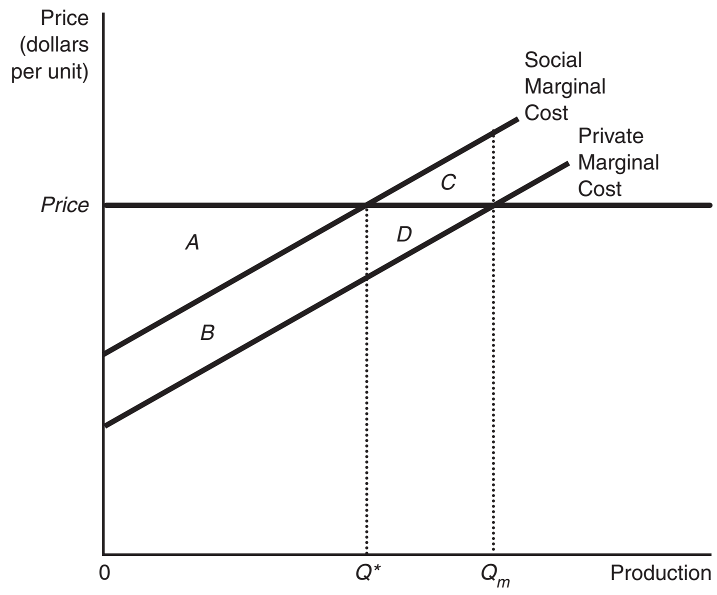
Private Resolution through Negotiation

The simplest means to restore efficiency occurs when the number of affected parties is small, making negotiation feasible. Suppose, for example, we return to the case used earlier in this chapter to illustrate an externality—the conflict between the polluting steel company and the downstream resort.

Figure 2.10 reveals the answer. If the resort offers a bribe of $C + D$, they would experience damage reduction from the decrease in production from Q_m to Q^*. Let's assume that the bribe is equal to this amount. Would the steel company be willing to reduce production to the desired level? If they refused the bribe, their producer surplus would be $A + B + D$. If they accepted the bribe, their producer surplus would be $A + B$ plus the bribe, so their total return would be $A + B + C + D$. Clearly, they are better off by C if they accept the bribe. Society as a whole is better off by amount C as well since the economic surplus from Q_m is $A - C$ and the economic surplus for Q^* is A.

Our discussion of individual negotiations raises two questions: (1) Should the property right always belong to the party who gained or seized it first (in this case the steel company)? (2) How can environmental risks be handled when prior negotiation is clearly impractical? These questions are routinely answered by the court system.

FIGURE 2.10 Efficient Output with Pollution Damage



The Courts: Property Rules and Liability Rules

The court system can respond to environmental conflicts by imposing either property rules or liability rules. Property rules specify the initial allocation of the entitlement. The entitlements at conflict in our example are, on one hand, the right to add waste products to the river and, on the other, the right to an attractive river. In applying property rules, the court merely decides which right is preeminent and places an injunction against violating that right. The injunction is removed only upon obtaining the consent of the party whose right was violated. Consent is usually obtained in return for an out-of-court monetary settlement.

Note that in the absence of a court decision, the entitlement is naturally allocated to the party that can most easily seize it. In our example, the natural allocation would give the entitlement to the steel company. The courts must decide whether to overturn this natural allocation.

How would they decide? And what difference would their decision make? The answers may surprise you.

In a classic article, economist Ronald Coase (1960) held that as long as negotiation costs are negligible and affected consumers can negotiate freely with each other (when the number of affected parties is small), the court could allocate the entitlement to *either* party, and an efficient allocation would result. The only effect of the court's decision would be to change the distribution of surplus among the affected parties. This remarkable conclusion has come to be known as the Coase theorem.

Why is this so? In Figure 2.10, we showed that if the steel company has the property right, it is in the resort's interest to offer a bribe that results in the desired level of output. Suppose, now, that the resort has the property right instead. To pollute in this case, the steel company must bribe the resort. Suppose it could pollute only if it compensated the resort for all damages. (In other words, it would agree to pay the difference between the two marginal cost curves up to the level of output actually chosen.) As long as this compensation was required, the steel company would choose to produce Q^* since that is the level at which its producer's surplus maximized. (Note that, due to the compensation, the curve the steel company uses to calculate its producer surplus is now the higher marginal cost curve.)

The difference between these different ways of allocating property rights lies in how the cost of obtaining the efficient level of output is shared between the parties. When the property right is assigned to the steel company, the cost is borne by the resort (part of the cost is the damage and part is the bribe to reduce the level of damage). When the property right is assigned to the resort, the cost is borne by the steel company (it now must compensate for all damage). In either case, the efficient level of production results. The Coase theorem shows that the very existence of an inefficiency triggers pressures for improvements. Furthermore, the existence of this pressure does not depend on the assignment of property rights.

This is an important point. As we shall see in succeeding chapters, private efforts triggered by inefficiency can frequently prevent the worst excesses of environmental degradation. Yet the importance of this theorem should not be overstated. Both theoretical and practical objections can be raised. The chief theoretical qualification concerns the assumption that wealth effects do not matter. The decision to confer the property right on a particular party results in a transfer of wealth to that party. This transfer might shift the demand curve for either steel or resorts out, as long as higher incomes result in greater demand. Whenever wealth effects are significant, the type of property rule issued by the court affects the outcome.

Wealth effects normally are small, so the zero-wealth-effect assumption is probably not a fatal flaw. Some serious practical flaws, however, do mar the usefulness of the Coase theorem. The first involves the incentives for polluting that result when the property right is assigned to the polluter. Since pollution would become a profitable activity with this assignment, other polluters might be encouraged to increase production and pollution in order to earn the bribes. That certainly would not be efficient.

Negotiation is also difficult to apply when the number of people affected by the pollution is large. You may have already noticed that in the presence of several affected parties, pollution reduction is a public good. The free-rider problem would make it difficult for the group to act cohesively and effectively for the restoration of efficiency.

When individual negotiation is not practical for one reason or another, the courts can turn to liability rules. These are rules that award monetary damages, after the fact, to the injured party. The amount of the award is designed to correspond to the amount of damage inflicted. Thus, returning to Figure 2.10, a liability rule would force the steel company to compensate the resort for all damages incurred. In this case, it could choose any production level it wanted, but it would have to pay the resort an amount of money equal to the area between the two marginal cost curves from the origin to the chosen level of output. In this case the steel plant would maximize its producer's surplus by choosing Q^*. (Why wouldn't the steel plant choose to produce more than that? Why wouldn't the steel plant choose to produce less than that?)

The moral of this story is that appropriately designed liability rules can also correct inefficiencies by forcing those who cause damage to bear the cost of that damage. Internalizing previously external costs causes profit-maximizing decisions to be compatible with efficiency.

Liability rules are interesting from an economics point of view because early decisions create precedents for later ones. Imagine, for example, how the incentives to prevent oil spills facing an oil company are transformed once it has a legal obligation to clean up after an oil spill and to compensate fishermen for reduced catches. It quickly becomes evident that in this situation accident prevention can become cheaper than retrospectively dealing with the damage once it has occurred.

This approach, however, also has its limitations. It relies on a case-by-case determination based on the unique circumstances for each case. Administratively, such a determination is very expensive. Expenses, such as court time, lawyers' fees, and so on, fall into a category called *transaction costs* by economists. In the present context, these are the administrative costs incurred in attempting to correct the inefficiency. When the number of parties involved in a dispute is large and the circumstances are common, we are tempted to correct the inefficiency by statutes or regulations rather than court decisions.

Legislative and Executive Regulation

These remedies can take several forms. The legislature could dictate that no one produce more steel or pollution than Q^*. This dictum might then be backed up with sufficiently large jail sentences or fines to deter potential violators. Alternatively, the legislature could impose a tax on steel or on pollution. A per-unit tax equal to the vertical distance between the two marginal cost curves would work (see Figure 2.10).

Legislatures could also establish rules to permit greater flexibility and yet reduce damage. For example, zoning laws might establish separate areas for steel plants and resorts. This approach assumes that the damage can be substantially reduced by keeping nonconforming uses apart.

They could also require the installation of particular pollution control equipment (as when catalytic converters were required on automobiles), or deny the use of a particular production ingredient (as when lead was removed from gasoline). In other words, they can regulate outputs, inputs, production processes, emissions, and even the location of production in their attempt to produce an efficient outcome. In subsequent chapters, we shall examine the various options policy-makers have not only to show how they can modify environmentally destructive behavior, but also to establish the degree to which they can promote efficiency.

Bribes are, of course, not the only means victims have at their disposal for lowering pollution. When the victims also consume the products produced by the polluters, consumer boycotts are possible. When the victims are employed by the polluter producer, strikes or other forms of labor resistance are possible.

An Efficient Role for Government

While the economic approach suggests that government action could well be used to restore efficiency, it also suggests that inefficiency is not a sufficient condition to justify government intervention. Any corrective mechanism involves transaction costs. If these transaction costs are high enough, and the surplus to be derived from correcting the inefficiency small enough, then it is best simply to live with the inefficiency.

Consider, for example, the pollution problem. Wood-burning stoves, which were widely used for cooking and heat in the late 1800s in the United States, were sources of pollution, but because of the enormous capacity of the air to absorb the emissions, no regulation resulted. More recently, however, the resurgence of demand for wood-burning stoves, precipitated in part by high oil prices, has resulted in strict regulations for wood-burning stove emissions because the population density is so much higher.

As society has evolved, the scale of economic activity and the resulting emissions have increased. Cities are experiencing severe problems from air and water pollutants because of the clustering of activities. Both the expansion and the clustering have increased the amount of emissions per unit volume of air or water. As a result, pollutant concentrations have caused perceptible problems with human health, vegetation growth, and aesthetics.

Historically, as incomes have risen, the demand for leisure activities has also risen. Many of these leisure activities, such as canoeing and backpacking, take place in unique, pristine environmental areas. With the number of these areas declining as a result of conversion to other uses, the value of remaining areas has increased. Thus, the value derived from protecting some areas have risen over time until they have exceeded the transaction costs of protecting them from pollution and/or development.

The level and concentration of economic activity, having increased pollution problems and driven up the demand for clean air and pristine areas, have created the preconditions for government action. Can government respond or will rent seeking prevent efficient political solutions? We devote much of this book to searching for the answer to that question.

Summary

How producers and consumers use the resources making up the environmental asset depends on the nature of the entitlements embodied in the property rights governing resource use. When property rights systems are exclusive, transferable, and enforceable, the owner of a resource has a powerful incentive to use that resource efficiently, since the failure to do so results in a personal loss.

The economic system will not always sustain efficient allocations, however. Specific circumstances that could lead to inefficient allocations include externalities; improperly defined property rights systems (such as open-access resources and public goods); and imperfect markets for trading the property rights to the resources (monopoly). When these circumstances arise, market allocations do not maximize the surplus.

Due to rent-seeking behavior by special interest groups or the less-than-perfect implementation of efficient plans, the political system can produce inefficiencies as well. Voter ignorance on many issues, coupled with the public-good nature of any results of political activity, tends to create a situation in which maximizing an individual's private surplus (through lobbying, for example) can be at the expense of a lower economic surplus for all consumers and producers.

The efficiency criterion can be used to assist in the identification of circumstances in which our political and economic institutions lead us astray. It can also assist in the search for remedies by facilitating the design of regulatory, judicial, or legislative solutions.

Discussion Questions

1. In a well-known legal case, *Miller v. Schoene* (287 U.S. 272), a classic conflict of property rights was featured. Red cedar trees, used only for ornamental purposes, carried a disease that could destroy apple orchards within a radius of two miles. There was no known way of curing the disease except by destroying the cedar trees or by ensuring that apple orchards were at least two miles away from the cedar trees. Apply the Coase theorem to this situation. Does it make any difference to the outcome whether the cedar tree owners are entitled to retain their trees or the apple growers are entitled to be free of them? Why or why not?
2. In primitive societies, the entitlements to use land were frequently possessory rights rather than ownership rights. Those on the land could use it as they wished, but they could not transfer it to anyone else. One could acquire a new plot by simply occupying and using it, leaving the old plot available for someone else. Would this type of entitlement system cause more or less incentive to conserve the land than an ownership entitlement? Why? Would a possessory entitlement system be more efficient in a modern society or a primitive society? Why?

Self-Test Exercises

1. Suppose the state is trying to decide how many miles of a very scenic river it should preserve. There are 100 people in the community, each of whom has an identical inverse demand function given by $P = 10 - 1.0q$, where q is the number of miles preserved and P is the per-mile price he or she is willing to pay for q miles of preserved river. (a) If the marginal cost of preservation is $500 per mile, how many miles would be preserved in an efficient allocation? (b) How large is the economic surplus?
2. Suppose the market demand function (expressed in dollars) for a normal product is $P = 80 - q$, and the marginal cost (in dollars) of producing it is $MC = 1q$, where P is the price of the product and q is the quantity demanded and/or supplied.
   a. How much would be supplied by a competitive market?
   b. Compute the consumer surplus and producer surplus. Show that their sum is maximized.
   c. Compute the consumer surplus and the producer surplus assuming this same product was supplied by a monopoly. (*Hint*: The marginal revenue curve has twice the slope of the demand curve.)
   d. Show that when this market is controlled by a monopoly, producer surplus is larger, consumer surplus is smaller, and the sum of the two surpluses is smaller than when the market is controlled by competitive industry.
3. Suppose you were asked to comment on a proposed policy to control oil spills. Since the average cost of an oil spill has been computed as $*X*, the proposed policy would require any firm responsible for a spill immediately to pay the government $*X*. Is this likely to result in the efficient amount of precaution against oil spills? Why or why not?
4. "In environmental liability cases, courts have some discretion regarding the magnitude of compensation polluters should be forced to pay for the environmental incidents they cause. In general, however, the larger the required payments the better." Discuss.
5. Label each of the following propositions as descriptive or normative and defend your choice:
   a. Energy efficiency programs would create jobs.
   b. Money spent on protecting endangered species is wasted.
   c. To survive our fisheries must be privatized.
   d. Raising transport costs lowers suburban land values.
   e. Birth control programs are counterproductive.
6. Identify whether each of the following resource categories is a public good, a common-pool resource, or neither and defend your answer:
   a. A pod of whales in the ocean to whale hunters.
   b. A pod of whales in the ocean to whale watchers.
   c. The benefits from reductions of greenhouse gas emissions.
   d. Water from a town well that excludes nonresidents.
   e. Bottled water.

Further Reading

Bromley, Daniel W. *Environment and Economy: Property Rights and Public Policy* (Oxford: Basil Blackwell, Inc., 1991). A detailed exploration of the property rights approach to environmental problems.

Bromley, Daniel W., ed. *Making the Commons Work: Theory, Practice and Policy* (San Francisco: ICS Press, 1992). An excellent collection of 13 essays exploring various formal and informal approaches to controlling the use of common-property resources.

Lueck, Dean. "The Extermination and Conservation of the American Bison," *Journal of Legal Studies* Vol 31(S2) (1989): s609–2652. A fascinating look at the role property rights played in the fate of the American bison.

Ostrom, Elinor, Thomas Dietz, Nives Dolsak, Paul Stern, Susan Stonich, and Elke U. Weber, eds. *The Drama of the Commons* (Washington, DC: National Academy Press, 2002). A compilation of articles and papers on common pool resources.

Ostrom, Elinor. *Crafting Institutions for Self-Governing Irrigation Systems* (San Francisco: ICS Press, 1992). Argues that common-pool problems are sometimes solved by voluntary organizations, rather than by a coercive state; among the cases considered are communal tenure in meadows and forests, irrigation communities, and fisheries.

Sandler, Todd. *Collective Action: Theory and Applications* (Ann Arbor: University of Michigan, 1992). A formal examination of the forces behind collective action's failures and successes.

Stavins, Robert N. "Harnessing Market Forces to Protect the Environment," *Environment* Vol. 31 (1989): 4–7, 28–35. An excellent, nontechnical review of the many ways in which the creative use of economic policies can produce superior environmental outcomes.

Stavins, Robert. *The Economics of the Environment: Selected Readings*, 5th ed. (New York: W.W. Norton and Company, 2005). A carefully selected collection of readings that would complement this text.

Additional References and Historically Significant References are available on this book's Companion Website:
http://www.pearsonhighered.com/tietenberg/

3 Evaluating Trade-Offs: Benefit–Cost Analysis and Other Decision-Making Metrics

No sensible decision can be made any longer without taking into account not only the world as it is, but the world as it will be . . .

—Isaac Asimov, US science fiction novelist and scholar (1920–1992)

Introduction

In the last chapter we noted that economic analysis has both positive and normative dimensions. The normative dimension helps to separate the policies that make sense from those that don't. Since resources are limited, it is not possible to undertake all ventures that might appear desirable so making choices is inevitable.

Normative analysis can be useful in public policy in several different situations. It might be used, for example, to evaluate the desirability of a proposed new pollution control regulation or a proposal to preserve an area currently scheduled for development. In these cases the analysis helps to provide guidance on the desirability of a program before that program is put into place. In other contexts it might be used to evaluate how an already-implemented program has worked out in practice. Here the relevant question is: Would this be (or was this) a wise use of resources? In this chapter, we present and demonstrate the use of several decision-making metrics that can assist us in evaluating options.

Normative Criteria for Decision Making

Normative choices can arise in two different contexts. In the first context we need simply to choose among options that have been predefined, while in the second we try to find the optimal choice among all the possible choices.

Evaluating Predefined Options: Benefit–Cost Analysis

If you were asked to evaluate the desirability of some proposed action, you would probably begin by attempting to identify both the gains and the losses from that action. If the gains exceed the losses, then it seems natural to support the action.

That simple framework provides the starting point for the normative approach to evaluating policy choices in economics. Economists suggest that actions have both benefits and costs. If the benefits exceed the costs, then the action is desirable. On the other hand, if the costs exceed the benefits, then the action is not desirable.

We can formalize this in the following way. Let B be the benefits from a proposed action and C be the costs. Our decision rule would then be

If $B > C$, support the action.

Otherwise, oppose the action.[1]

As long as B and C are positive, a mathematically equivalent formulation would be

If $B/C > 1$, support the action.

Otherwise, oppose the action.

So far so good, but how do we measure benefits and costs? In economics the system of measurement is anthropocentric, which simply means human centered. All benefits and costs are valued in terms of their effects (broadly defined) on humanity. As shall be pointed out later, that does *not* imply (as it might first appear) that ecosystem effects are ignored unless they *directly* affect humans. The fact that large numbers of humans contribute voluntarily to organizations that are dedicated to environmental protection provides ample evidence that humans place a value on environmental preservation that goes well beyond any direct use they might make of it. Nonetheless, the notion that humans are doing the valuing is a controversial point that will be revisited and discussed in Chapter 4 along with the specific techniques for valuing these effects.

In benefit–cost analysis, benefits are measured simply as the relevant area under the demand curve since the demand curve reflects consumers' willingness to pay. Total costs are measured by the relevant area under the marginal cost curve.

It is important to stress that environmental services have costs even though they are produced without any human input. All costs should be measured as opportunity costs. As presented in Example 3.1, the *opportunity cost* for using resources in a new or an alternative way is the net benefit lost when specific environmental services are foregone in the conversion to the new use. The notion that it is costless to convert a forest to a new use is obviously wrong if valuable ecological or human services are lost in the process.

To firm up this notion of opportunity cost, consider another example. Suppose a particular stretch of river can be used either for white-water canoeing or to generate electric power. Since the dam that generates the power would flood the rapids, the two uses are incompatible. The opportunity cost of producing power is the foregone net benefit that would have resulted from the white-water canoeing. The *marginal opportunity cost curve* defines the additional cost of producing another unit of electricity resulting from the associated incremental loss of net benefits due to reduced opportunities for white-water canoeing.

---

[1]Actually if $B = C$, it wouldn't make any difference if the action occurs or not; the benefits and costs are a wash.

EXAMPLE 3.1

Valuing Ecological Services from Preserved Tropical Forests

As Chapter 12 makes clear, one of the main threats to tropical forests is the conversion of forested land to some other use (agriculture, residences, and so on). Whether economic incentives favor conversion of the land depends upon the magnitude of the value that would be lost through conversion. How large is that value? Is it large enough to support preservation?

A group of ecologists investigated this question for a specific set of tropical forest fragments in Costa Rica. They chose to value one specific ecological service provided by the local forest: wild bees using the nearby tropical forest as a habitat provided pollination services to aid coffee production. While this coffee (*C. arabica*) can self-pollinate, pollination from wild bees has been shown to increase coffee productivity from 15 to 50 percent.

When the authors placed an economic value on this particular ecological service, they found that the pollination services from two specific preserved forest fragments (46 and 111 hectares, respectively) were worth approximately $60,000 per year for one large, nearby Costa Rican coffee farm. As the authors conclude:

> *The value of forest in providing crop pollination service alone is . . . of at least the same order [of magnitude] as major competing land uses, and infinitely greater than that recognized by most governments (i.e., zero).*

These estimates only partially capture the value of this forest because they consider only a single farm and a single type of ecological service. (This forest also provides carbon storage and water purification services, for example, and these were not included in the calculation.) Despite their partial nature, however, these calculations already begin to demonstrate the economic value of preserving the forest, even when considering only a limited number of specific instrumental values.

Source: Taylor H. Ricketts et al., "Economic Value of Tropical Forest to Coffee Production." PNAS (Proceedings of the National Academy of Science), Vol. 101, No. 34, August 24, 2002, pp. 12579–12582.

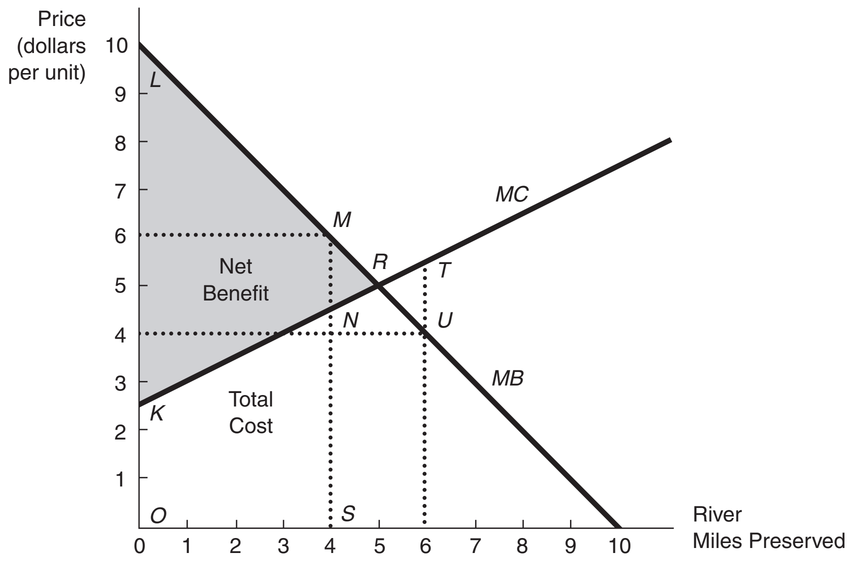
Since net benefit is defined as the excess of benefits over costs, it follows that net benefit is equal to that portion of the area under the demand curve that lies above the supply curve.

Consider Figure 3.1, which illustrates the net benefits from preserving a stretch of river. Let's use this example to illustrate the use of the decision rules introduced earlier. For example, let's suppose that we are considering preserving a four-mile stretch of river and that the benefits and costs of that action are reflected in Figure 3.1. Should that stretch be preserved? Why or why not? We will return to this example later.

Finding the Optimal Outcome

In the preceding section we examined how benefit–cost analysis can be used to evaluate the desirability of specific actions. In this section we want to examine how this approach can be used to identify "optimal" or best approaches.

FIGURE 3.1 The Derivation of Net Benefits



In subsequent chapters, which address individual environmental problems, the normative analysis will proceed in three steps. First we will identify an optimal outcome. Second we will attempt to discern the extent to which our institutions produce optimal outcomes and, where divergences occur between actual and optimal outcomes, to attempt to uncover the behavioral sources of the problems. Finally we can use both our knowledge of the nature of the problems and their underlying behavioral causes as a basis for designing appropriate policy solutions. Although applying these three steps to each of the environmental problems must reflect the uniqueness of each situation, the overarching framework used to shape that analysis is the same.

To provide some illustrations of how this approach is used in practice, consider two examples: one drawn from natural resource economics and another from environmental economics. These are meant to be illustrative and to convey a flavor of the argument; the details are left to upcoming chapters.

Consider the rising number of depleted ocean fisheries. Depleted fisheries, which involve fish populations that have fallen so low as to threaten their viability as commercial fisheries, not only jeopardize oceanic biodiversity, but also pose a threat to both the individuals who make their living from the sea and the communities that depend on fishing to support their local economies.

How would an economist attempt to understand and resolve this problem? The first step would involve defining the optimal stock or the optimal rate of harvest of the fishery. The second step would compare this level with the actual stock and harvest levels. Once this economic framework is applied, not only does it become clear that stocks are much lower than optimal for many fisheries, but also the reason for excessive exploitation becomes clear. Understanding the nature of the problem has led quite naturally to some solutions. Once implemented, these policies have allowed some fisheries to begin the process of renewal. The details of this analysis and the policy implications that flow from it are covered in Chapter 13.

Another problem involves solid waste. As local communities run out of room for landfills in the face of an increasing generation of waste, what can be done?

Economists start by thinking about how one would define the optimal amount of waste. The definition necessarily incorporates waste reduction and recycling as aspects of the optimal outcome. The analysis not only reveals that current waste levels are excessive, but also suggests some specific behavioral sources of the problem. Based upon this understanding, specific economic solutions have been identified and implemented. Communities that have adopted these measures have generally experienced lower levels of waste and higher levels of recycling. The details are spelled out in Chapter 8.

In the rest of the book, similar analysis is applied to population, energy, minerals, agriculture, air and water pollution, and a host of other topics. In each case the economic analysis helps to point the way toward solutions. To initiate that process we must begin by defining "optimal."

Relating Optimality to Efficiency

According to the normative choice criterion introduced earlier in this chapter, desirable outcomes are those where the benefits exceed the costs. It is therefore a logical next step to suggest that optimal polices are those that maximize net benefits (benefits–costs). The concept of *static efficiency*, or merely *efficiency*, was introduced in Chapter 2. An allocation of resources is said to satisfy the static efficiency criterion if the economic surplus from the use of those resources is maximized by that allocation. Notice that the net benefits area to be maximized in an "optimal outcome" for public policy is identical to the "economic surplus" that is maximized in an efficient allocation. Hence efficient outcomes are also optimal outcomes.

Let's take a moment to show how this concept can be applied. Previously we asked whether an action that preserved four miles of river was worth doing (Figure 3.1). The answer was yes because the net benefits from that action were positive.

Static efficiency, however, requires us to ask a rather different question, namely, what is the optimal (or efficient) number of miles to be preserved? We know from the definition that the optimal amount of preservation would maximize net benefits. Does preserving four miles maximize net benefits? Is it the efficient outcome?

We can answer that question by establishing whether it is possible to increase the net benefit by preserving more or less of the river. If the net benefit can be increased by preserving more miles, clearly, preserving four miles could not have maximized the net benefit and, therefore, could not have been efficient.

Consider what would happen if society were to choose to preserve five miles instead of four. Refer back to Figure 3.1. What happens to the net benefit? It increases by area *MNR*. Since we can find another allocation with greater net benefit, four miles of preservation could not have been efficient. Could five? Yes. Let's see why.

We know that five miles of preservation convey more net benefits than four. If this allocation is efficient, then it must also be true that the net benefit is smaller for levels of preservation higher than five. Notice that the additional cost of preserving the sixth unit (the area under the marginal cost curve) is larger than the additional benefit received from preserving it (the corresponding area under the demand curve). Therefore, the triangle *RTU* represents the reduction in net benefit that occurs if six miles are preserved rather than five.

Since the net benefit is reduced, both by preserving less than five and by preserving more than five, we conclude that five units is the preservation level that maximizes net benefit (the shaded area). Therefore, from our definition, preserving five miles constitutes an efficient or optimal allocation.[2]

One implication of this example, which will be very useful in succeeding chapters, is what we shall call the "first equimarginal principle":

First Equimarginal Principle (the "Efficiency Equimarginal Principle"): Social net benefits are maximized when the social marginal benefits from an allocation equal the social marginal costs.

The social marginal benefit is the increase in social benefits received from supplying one more unit of the good or service, while social marginal cost is the increase in cost incurred from supplying one more unit of the good or service.

This criterion helps to minimize wasted resources, but is it fair? The ethical basis for this criterion is derived from a concept called *Pareto optimality*, named after the Italian-born Swiss economist Vilfredo Pareto, who first proposed it around the turn of the twentieth century.

Allocations are said to be Pareto optimal if no other feasible allocation could benefit at least one person without any deleterious effects on some other person.

Allocations that do not satisfy this definition are suboptimal. Suboptimal allocations can always be rearranged so that some people can gain net benefits without the rearrangement causing anyone else to lose net benefits. Therefore, the gainers could use a portion of their gains to compensate the losers sufficiently to ensure they were at least as well off as they were prior to the reallocation.

Efficient allocations are Pareto optimal. Since net benefits are maximized by an efficient allocation, it is not possible to increase the net benefit by rearranging the allocation. Without an increase in the net benefit, it is impossible for the gainers to compensate the losers sufficiently; the gains to the gainers would necessarily be smaller than the losses to the losers.

Inefficient allocations are judged inferior because they do not maximize the size of the pie to be distributed. By failing to maximize net benefit, they are forgoing an opportunity to make some people better off without harming others.

---

[2]The monetary worth of the net benefit is the sum of two right triangles, and it equals (1/2)($5)(5) + (1/2)($2.50)(5) or $18.75. Can you see why?

Comparing Benefits and Costs Across Time

The analysis we have covered so far is very useful for thinking about actions where time is not an important factor. Yet many of the decisions made now have consequences that persist well into the future. Time is a factor. Exhaustible energy resources, once used, are gone. Biological renewable resources (such as fisheries or forests) can be overharvested, leaving smaller and possibly weaker populations for future generations. Persistent pollutants can accumulate over time. How can we make choices when the benefits and costs may occur at different points in time?

Incorporating time into the analysis requires an extension of the concepts we have already developed. This extension provides a way for thinking not only about the magnitude of benefits and costs, but also about their timing. In order to incorporate timing, the decision rule must provide a way to compare net benefits received in different time periods. The concept that allows this comparison is called *present value*. Therefore, before introducing this expanded decision rule, we must define present value.

Present value explicitly incorporates the time value of money. A dollar today invested at 10 percent interest yields $1.10 a year from now (the return of the $1 principal plus $0.10 interest). The present value of $1.10 received one year from now is therefore $1, because given $1 now, you can turn it into $1.10 a year from now by investing it at 10 percent interest. We can find the present value of any amount of money (X) received one year from now by computing $X/(1 + r)$, where r is the appropriate interest rate (10 percent in our above example).

What could your dollar earn in two years at r percent interest? Because of compound interest, the amount would be $\$1(1 + r)(1 + r) = \$1(1 + r)^2$. It follows then that the present value of X received two years from now is $X/(1 + r)^2$.

By now the pattern should be clear. The present value of a *one-time* net benefit received n years from now is

$$PV[B_n] = \frac{B_n}{(1 + r)^n}$$

The present value of a stream of net benefits $\{B_0, \ldots, B_n\}$ received over a period of n years is computed as

$$PV[B_0, \ldots, B_n] = \sum_{i=0}^{n} \frac{B_i}{(1 + r)^i}$$

where r is the appropriate interest rate and B_0 is the amount of net benefits received immediately. The process of calculating the present value is called *discounting*, and the rate r is referred to as the discount rate.

The number resulting from a present-value calculation has a straightforward interpretation. Suppose you were investigating an allocation that would yield the following pattern of net benefits on the last day of each of the next five years: $3,000,

TABLE 3.1 Demonstrating Present Value Calculations

Year	1	2	3	4	5	Sum
Annual Amounts	\$3,000	\$5,000	\$6,000	\$10,000	\$12,000	\$36,000
Present Value (r = 0.06)	\$2,830.19	\$4,449.98	\$5,037.72	\$7,920.94	\$8,967.10	\$29,205.92

TABLE 3.2 Interpreting Present Value Calculations

Year	1	2	3	4	5	6
Balance at Beginning of Year	\$29,205.92	\$27,958.28	\$24,635.77	\$20,113.92	\$11,320.75	\$0.00
Year-End Fund Balance before Payment (r = 0.06)	\$30,958.28	\$29,635.77	\$26,113.92	\$21,320.75	\$12,000.00	
Payment	\$3,000	\$5,000	\$6,000	\$10,000	\$12,000	

\$5,000, \$6,000, \$10,000, and \$12,000. If you use an interest rate of 6 percent (r = 0.06) and the above formula, you will discover that this stream has a present value of \$29,205.92 (see Table 3.1). Notice how each amount is discounted back the appropriate number of years to the present and then these discounted values are summed.

What does that number mean? If you put \$29,205.92 in a savings account earning 6 percent interest and wrote yourself checks, respectively, for \$3,000, \$5,000, \$6,000, \$10,000, and \$12,000 on the last day of each of the next five years, your last check would just restore the account to a \$0 balance (see Table 3.2). Thus, you should be indifferent about receiving \$29,205.92 now or in the specific five-year stream of benefits totaling \$36,000; given one, you can get the other. Hence, the method is called present value because it translates everything back to its current worth.

It is now possible to show how this analysis can be used to evaluate actions. Calculate the present value of net benefits from the action. If the present value is greater than zero, the action should be supported. Otherwise it should not.

Dynamic Efficiency

The static efficiency criterion is very useful for comparing resource allocations when time is not an important factor. How can we think about optimal choices when the benefits and costs occur at different points in time?

The traditional criterion used to find an optimal allocation when time is involved is called *dynamic efficiency*, a generalization of the static efficiency concept already developed. In this generalization, the present-value criterion provides a way for comparing the net benefits received in one period with the net benefits received in another.

An allocation of resources across n time periods satisfies the dynamic efficiency criterion if it maximizes the present value of net benefits that could be received from all the possible ways of allocating those resources over the n periods.

Applying the Concepts

Having now spent some time developing the concepts we need, let's take a moment to examine some actual studies in which they have been used.

Pollution Control

Benefit–cost analysis has been used to assess the desirability of efforts to control pollution. Pollution control certainly confers many benefits, but it also has costs. Do the benefits justify the costs? That was a question the U.S. Congress wanted answered, so in Section 812 of the Clean Air Act Amendments of 1990 it required the U.S. Environmental Protection Agency (EPA) to evaluate the benefits and costs of the U.S. air pollution control policy initially over the 1970–1990 period and subsequently over the 1990–2020 time period (see Example 3.2).

EXAMPLE 3.2

Does Reducing Pollution Make Economic Sense? Evidence from the Clean Air Act

In its 1997 report to Congress, the EPA presented the results of its attempt to discover whether the Clean Air Act had produced positive net benefits over the period 1970–1990. The results suggested that the present value of benefits (using a discount rate of 5 percent) was $22.2 trillion, while the costs were $0.523 trillion. Performing the necessary subtraction reveals that the net benefits were therefore equal to $21.7 trillion. According to this study, U.S. air pollution control policy during this period made very good economic sense.

Soon after the period covered by this analysis, substantive changes were made in the Clean Air Act Amendments of 1990 (the details of those changes are covered in later chapters). Did those additions also make economic sense?

In August of 2010, the U.S. EPA issued a report of the benefits and costs of the Clean Air Act from 1990 to 2020. This report suggests that the costs of meeting the 1990 Clean Air Act Amendment requirements are expected to rise to approximately $65 billion per year by 2020 (2006 dollars). Almost half of the compliance costs ($28 billion) arise from pollution controls placed on cars, trucks, and buses, while another $10 billion arises from reducing air pollution from electric utilities.

These actions are estimated to cause benefits (from reduced pollution damage) to rise from roughly $800 billion in 2000 to almost $1.3 trillion in 2010, ultimately reaching approximately $2 trillion per year (2006 dollars) by 2020! For persons living in the United States, a cost of approximately $200 per person by 2020 produces approximately a $6,000 gain in benefits from the improvement in air quality. Many of the estimated benefits come from reduced risk of early mortality due to exposure to fine particulate matter. Table 3.3 provides a summary of the costs and benefits and includes a calculation of the benefit/cost ratio.

TABLE 3.3 Summary Comparison of Benefits and Costs from the Clean Air Act-1990–2020 (Estimates in Million 2006$)

	Annual Estimates			Present Value Estimate
	2000	2010	2020	1990–2020
Monetized Direct Costs:				
Low[1]				
Central	$20,000	$53,000	$65,000	$380,000
High[1]				
Monetized Direct Benefits:				
Low[2]	$90,000	$160,000	$250,000	$1,400,000
Central	$770,000	$1,300,000	$2,000,000	$12,000,000
High[2]	$2,300,000	$3,800,000	$5,700,000	$35,000,000
Net Benefits:				
Low	$70,000	$110,000	$190,000	$1,000,000
Central	$750,000	$1,200,000	$1,900,000	$12,000,000
High	$2,300,000	$3,700,000	$5,600,000	$35,000,000
Benefit/Cost Ratio:				
Low[3]	5/1	3/1	4/1	4/1
Central	39/1	25/1	31/1	32/1
High[3]	115/1	72/1	88/1	92/1

[1]The cost estimates for this analysis are based on assumptions about future changes in factors such as consumption patterns, input costs, and technological innovation. We recognize that these assumptions introduce significant uncertainty into the cost results; however, the degree of uncertainty or bias associated with many of the key factors cannot be reliably quantified. Thus, we are unable to present specific low and high cost estimates.

[2]Low and high benefit estimates are based on primary results and correspond to 5th and 95th percentile results from statistical uncertainty analysis, incorporating uncertainties in physical effects and valuation steps of benefits analysis. Other significant sources of uncertainty not reflected include the value of unquantified or unmonetized benefits that are not captured in the primary estimates and uncertainties in emissions and air quality modeling.

[3]The low benefit/cost ratio reflects the ratio of the low benefits estimate to the central costs estimate, while the high ratio reflects the ratio of the high benefits estimate to the central costs estimate. Because we were unable to reliably quantify the uncertainty in cost estimates, we present the low estimate as "less than X," and the high estimate as "more than Y," where X and Y are the low and high benefit/cost ratios, respectively.

Sources: U.S. Environmental Protection Agency, THE BENEFITS AND COSTS OF THE CLEAN AIR ACT, 1970 to 1990 (Washington, DC: Environmental Protection Agency, 1997), Table 18, p. 56;. and the U.S. Environmental Protection Agency Office of Air and Radiation, THE BENEFITS AND COSTS OF THE CLEAN AIR ACT, 1990 to 2020 – Summary Report, 8/16/2010 and Full Report available at http://www.epa.gov/oar/sect812/prospective2.html (accessed on 12/31/2010).

In responding to this congressional mandate, the EPA set out to quantify and monetize the benefits and costs of achieving the emissions reductions required by U.S. policy. Benefits quantified by this study included reduced death rates and lower incidences of chronic bronchitis, lead poisoning, strokes, respiratory diseases, and heart disease as well as the benefits of better visibility, reduced structural damages, and improved agricultural productivity.

We shall return to this study later in the book for a deeper look at how these estimates were derived, but a couple of comments are relevant now. First, despite the fact that this study did not attempt to value all pollution damage to ecosystems that was avoided by this policy, the net benefits are still strongly positive. While presumably the case for controlling pollution would have been even stronger had all such avoided damage been included, the desirability of this form of control is evident even with only a partial consideration of benefits. An inability to monetize everything does not necessarily jeopardize the ability to reach sound policy conclusions.

Although these results justify the conclusion that pollution control made economic sense, they do not justify the stronger conclusion that the policy was efficient. To justify that conclusion, the study would have had to show that the present value of net benefits was maximized, not merely positive. In fact, this study did not attempt to calculate the maximum net benefits outcome and if it had, it would have almost certainly discovered that the policy during this period was not optimal. As we shall see in Chapters 15 and 16, the costs of the chosen policy approach were higher than necessary to achieve the desired emissions reductions. With an optimal policy mix, the net benefits would have been even higher.

Preservation versus Development

One of the most basic conflicts faced by environmental policy occurs when a currently underdeveloped but ecologically significant piece of land becomes a candidate for development. If developed, the land may not only provide jobs for workers, wealth for owners, and goods for consumers, but also it may degrade the ecosystem, possibly irreversibly. Wildlife habitat may be eliminated, wetlands may be paved over, and recreational opportunities may be gone forever. On the other hand, if the land were preserved, the specific ecosystem damages caused by development could be prevented, but the opportunity for increased income and employment provided by development would have been lost. These conflicts become intensified if unemployment rates in the area are high and the local ecology is rather unique.

One such conflict arose in Australia from a proposal to mine a piece of land in an area known as the Kakadu Conservation Zone (KCZ). Decision makers at that time had to decide whether it should be mined or preserved. One way to examine that question is to use the techniques above to examine the net benefits of the two alternatives (see Example 3.3).

EXAMPLE 3.3

Choosing between Preservation and Development in Australia

The Kakadu Conservation Zone, a 50-square-kilometer area lying entirely within the Kakadu National Park (KNP), was initially set aside by the government as part of a grazing lease. The current issue was whether it should be mined (it was believed to contain significant deposits of gold, platinum, and palladium) or added to the KNP, one of Australia's major parks. In recognition of its unique ecosystem and extensive wildlife as well as its aboriginal archeological sites, much of the park has been placed on the U.N. World Heritage List.

Mining would produce income and employment, but it could also cause the ecosystems in both the KCZ and KNP to experience irreversible damage. What value was to be placed on those risks? Would those risks outweigh the employment and income effects from mining?

To provide answers to these crucial questions, economists conducted a benefit–cost analysis using a technique known as contingent valuation. (We shall go into some detail about how this technique works in Chapter 4, but for now it can suffice to note that this is a technique for eliciting "willingness-to-pay" information.) The value of preserving the site was estimated to be A$435 million, while the present value of mining the site was estimated to be A$102 million.

According to this analysis, preservation was the preferred option and it was the option chosen by the government.

Source: Richard T. Carson, Leanne Wilks, and David Imber, "Valuing the Preservation of Australia's Kakadu Conservation Zone." OXFORD ECONOMIC PAPERS, Vol. 46, Supplement (1994), pp. 727–749.

Issues in Benefit Estimation

The analyst charged with the responsibility for performing a benefit–cost analysis encounters many decision points requiring judgment. If we are to understand benefit–cost analysis, the nature of these judgments must be clear in our minds.

Primary versus Secondary Effects. Environmental projects usually trigger both primary and secondary consequences. For example, the primary effect of cleaning a lake will be an increase in recreational uses of the lake. This primary effect will cause a further ripple effect on services provided to the increased number of users of the lake. Are these secondary benefits to be counted?

The answer depends upon the employment conditions in the surrounding area. If this increase in demand results in employment of previously unused resources, such as labor, the value of the increased employment should be counted. If, on the other hand, the increase in demand is met by a shift in previously employed resources from one use to another, it is a different story. In general, secondary employment benefits should be counted in high unemployment areas or when the particular skills demanded are underemployed at the time the project is commenced.

This should not be counted when the project simply results in a rearrangement of productively employed resources.

Accounting Stance. The accounting stance refers to the geographic scale at which the benefits are measured. Who benefits? If a proposed project is funded by a national government, but benefits a local or regional area, a benefit-cost analysis will look quite different depending on whether the analysis is done at the regional or national scale.

With and Without Principle. The "with and without" principle states that only those benefits that would result from the project should be counted, ignoring those that would have accrued anyway. Mistakenly including benefits that would have accrued anyway would overstate the benefits of the program.

Tangible versus Intangible Benefits. *Tangible benefits* are those that can reasonably be assigned a monetary value. *Intangible benefits* are those that cannot be assigned a monetary value, either because data are not available or reliable enough or because it is not clear how to measure the value even with data.[3] Quantification of intangible benefits is the primary topic of the next chapter.

How are intangible benefits to be handled? One answer is perfectly clear: They should not be ignored. To ignore intangible benefits is to bias the results. That benefits are intangible does not mean they are unimportant.

Intangible benefits should be quantified to the fullest extent possible. One frequently used technique is to conduct a sensitivity analysis of the estimated benefit values derived from less than perfectly reliable data. We can determine, for example, whether or not the outcome is sensitive, within wide ranges, to the value of this benefit. If not, then very little time has to be spent on the problem. If the outcome is sensitive, the person or persons making the decision bear the ultimate responsibility for weighing the importance of that benefit.

Approaches to Cost Estimation

Estimating costs is generally easier than estimating benefits, but it is not easy. One major problem for both derives from the fact that benefit–cost analysis is forward-looking and thus requires an estimate of what a particular strategy *will* cost, which is much more difficult than tracking down what an existing strategy *does* cost.

Two approaches have been developed to estimate these costs.

The Survey Approach. One way to discover the costs associated with a policy is to ask those who bear the costs, and presumably know the most about them, to reveal the magnitude of the costs to policy-makers. Polluters, for example, could be

[3]The division between tangible and intangible benefits changes as our techniques improve. Recreation benefits were, until the advent of the travel-cost model, treated as intangible. The travel cost model will be discussed in the next chapter.

asked to provide control-cost estimates to regulatory bodies. The problem with this approach is the strong incentive not to be truthful. An overestimate of the costs can trigger less stringent regulation; therefore, it is financially advantageous to provide overinflated estimates.

The Engineering Approach. The engineering approach bypasses the source being regulated by using general engineering information to catalog the possible technologies that could be used to meet the objective and to estimate the costs of purchasing and using those technologies. The final step in the engineering approach is to assume that the sources would use technologies that minimize cost. This produces a cost estimate for a "typical," well-informed firm.

The engineering approach has its own problems. These estimates may not approximate the actual cost of any particular firm. Unique circumstances may cause the costs of that firm to be higher, or lower, than estimated; the firm, in short, may not be typical.

The Combined Approach. To circumvent these problems, analysts frequently use a combination of survey and engineering approaches. The survey approach collects information on possible technologies, as well as special circumstances facing the firm. Engineering approaches are used to derive the actual costs of those technologies, given the special circumstances. This combined approach attempts to balance information best supplied by the source with that best derived independently.

In the cases described so far, the costs are relatively easy to quantify and the problem is simply finding a way to acquire the best information. This is not always the case, however. Some costs are not easy to quantify, although economists have developed some ingenious ways to secure monetary estimates even for those costs.

Take, for example, a policy designed to conserve energy by forcing more people to carpool. If the effect of this is simply to increase the average time of travel, how is this cost to be measured?

For some time, transportation analysts have recognized that people value their time, and quite a literature has now evolved to provide estimates of how valuable time savings or time increases would be. The basis for this valuation is opportunity cost—how the time might be used if it weren't being consumed in travel. Although the results of these studies depend on the amount of time involved, individuals seem to value their travel time at a rate not more than half their wage rates.

The Treatment of Risk

For many environmental problems, it is not possible to state with certainty what consequences a particular policy will have, because scientific estimates themselves often are imprecise. Determining the efficient exposure to potentially toxic substances requires obtaining results at high doses and extrapolating to low doses, as well as extrapolating from animal studies to humans. It also requires relying upon epidemiological studies that infer a pollution-induced adverse human health impact from correlations between indicators of health in human populations and recorded pollution levels.

For example, consider the potential damages from climate change. While most scientists now agree on the potential impacts of climate change, such as sea level rise and species losses, the timing and extent of those losses are not certain.

The treatment of risk in the policy process involves two major dimensions: (1) identifying and quantifying the risks; and (2) deciding how much risk is acceptable. The former is primarily scientific and descriptive, while the latter is more evaluative or normative.

Benefit–cost analysis grapples with the evaluation of risk in several ways. Suppose we have a range of policy options *A*, *B*, *C*, *D* and a range of possible outcomes *E*, *F*, *G* for each of these policies depending on how the economy evolves over the future. These outcomes, for example, might depend on whether the demand growth for the resource is low, medium, or high. Thus, if we choose policy *A*, we might end up with outcomes *AE*, *AF*, or *AG*. Each of the other policies has three possible outcomes as well, yielding a total of 12 possible outcomes.

We could conduct a separate benefit–cost analysis for each of the 12 possible outcomes. Unfortunately, the policy that maximizes net benefits for *E* may be different from that which maximizes net benefits for *F* or *G*. Thus, if we only knew which outcome would prevail, we could select the policy that maximized net benefits; the problem is that we do not. Furthermore, choosing the policy that is best if outcome *E* prevails may be disastrous if *G* results instead.

When a dominant policy emerges, this problem is avoided. A *dominant policy* is one that confers higher net benefits for every outcome. In this case, the existence of risk concerning the future is not relevant for the policy choice. This fortuitous circumstance is exceptional rather than common, but it can occur.

Other options exist even when dominant solutions do not emerge. Suppose, for example, that we were able to assess the likelihood that each of the three possible outcomes would occur. Thus, we might expect outcome *E* to occur with probability 0.5, *F* with probability 0.3, and *G* with probability 0.2. Armed with this information, we can estimate the expected present value of net benefits. The *expected present value of net benefits* for a particular policy is defined as the sum over outcomes of the present value of net benefits for that policy where each outcome is weighted by its probability of occurrence. Symbolically this is expressed as

$$EPVNB_j = \sum_{i=0}^{I} P_i PVNB_{ij}, \quad j = 1, \ldots, J, \tag{3.1}$$

where

$EPVNB_j$ = expected present value of net benefits for policy j

P_i = probability of the ith outcome occurring

$PVNB_{ij}$ = present value of net benefits for policy j if outcome i prevails

J = number of policies being considered

I = number of outcomes being considered

The final step is to select the policy with the highest expected present value of net benefits.

This approach has the substantial virtue that it weighs higher probability outcomes more heavily. It also, however, makes a specific assumption about society's preference for risk. This approach is appropriate if society is risk-neutral. *Risk-neutrality* can be defined most easily by the use of an example. Suppose you were allowed to choose between being given a definite $50 or entering a lottery in which you had a 50 percent chance of winning $100 and a 50 percent chance of winning nothing. (Notice that the expected value of this lottery is $50 = 0.5($100) + 0.5($0).) You would be said to be risk-neutral if you would be indifferent between these two choices. If you view the lottery as more attractive, you would be exhibiting *risk-loving* behavior, while a preference for the definite $50 would suggest *risk-averse* behavior. Using the expected present value of net benefits approach implies that society is risk-neutral.

Is that a valid assumption? The evidence is mixed. The existence of gambling suggests that at least some members of society are risk-loving, while the existence of insurance suggests that, at least for some risks, others are risk-averse. Since the same people may gamble and own insurance policies, it is likely that the type of risk may be important.

Even if individuals were demonstrably risk-averse, this would not be a sufficient condition for the government to forsake risk-neutrality in evaluating public investments. One famous article (Arrow and Lind, 1970) argues that risk-neutrality is appropriate since "when the risks of a public investment are publicly borne, the total cost of risk-bearing is insignificant and, therefore, the government should ignore uncertainty in evaluating public investments." The logic behind this result suggests that as the number of risk bearers (and the degree of diversification of risks) increases, the amount of risk borne by any individual diminishes to zero.

When the decision is irreversible, as demonstrated by Arrow and Fisher (1974), considerably more caution is appropriate. Irreversible decisions may subsequently be regretted, but the option to change course will be lost forever. Extra caution also affords an opportunity to learn more about alternatives to this decision and its consequences before acting. Isn't it comforting to know that occasionally procrastination can be optimal?

There is a movement in national policy in both the courts and the legislature to search for imaginative ways to define acceptable risk. In general, the policy approaches reflect a case-by-case method. We shall see that current policy reflects a high degree of risk aversion toward a number of environmental problems.

Distribution of Benefits and Costs

Many agencies are now required to consider the distributional impacts of costs and benefits as part of any economic analysis. For example, the U.S. EPA provides guidelines on distributional issues in its "Guidelines for Preparing Economic Analysis." According to the EPA, distributional analysis "assesses changes in social welfare by examining the effects of a regulation across different sub-populations and entities." Distributional analysis can take two forms: *economic impact analysis* and *equity analysis*. Economic impact analysis focuses on a broad characterization of who gains and who loses from a given policy. Equity analysis examines impacts on disadvantaged groups or sub-populations. The latter delves into the normative issue of equity or fairness in the distribution of costs and benefits. The issue of environmental justice will be considered further in Chapter 19.[4] Loomis (2011) outlines several approaches for incorporating distribution and equity into benefit–cost analysis.

Choosing the Discount Rate

The discount rate can be defined conceptually as the social opportunity cost of capital. This cost of capital can be divided further into two components: (1) the riskless cost of capital and (2) the risk premium. The choice of the discount rate can influence policy decisions. Recall that discounting allows us to compare all costs and benefits in current dollars, regardless of when the benefits accrue or costs are charged. Suppose, a project will impose an immediate cost of $4,000,000 (today's dollars), but the $5,500,000 benefits will not be earned until 5 years out. Is this project a good idea? On the surface it might seem like it is, but recall that $5,500,000 in 5 years is not the same as $5,500,000 today. At a discount rate of 5 percent, the present value of benefits minus the present value of costs is positive. However, at a 10 percent discount rate, this same calculation yields a negative value, since the present value of costs exceeds the benefits. Can you reproduce the calculations that yield these conclusions?

As Example 3.4 indicates, this has been, and continues to be, an important issue. When the public sector uses a discount rate lower than that in the private sector, the public sector will find more projects with longer payoff periods worthy of authorization. And, as we have already seen, the discount rate is a major determinant of the allocation of resources among generations as well.

Traditionally, economists have used long-term interest rates on government bonds as one measure of the cost of capital, adjusted by a risk premium that would depend on the riskiness of the project considered. Unfortunately, the choice of how large an adjustment to make has been left to the discretion of the analysts. This ability to affect the desirability of a particular project or policy by the choice of discount rate led to a situation in which government agencies were using a variety of discount rates to justify programs or projects they supported. One set of hearings conducted by Congress during the 1960s discovered that, at one time, agencies were using discount rates ranging from 0 to 20 percent.

During the early 1970s the Office of Management and Budget published a circular that required, with some exceptions, all government agencies to use a discount rate of 10 percent in their benefit–cost analysis. A revision issued in 1992 reduced the required discount rate to 7 percent. This circular also includes guidelines for benefit–cost analysis and specifies that certain rates will change annually.[5] This standardization reduces biases by eliminating the agency's ability to choose a discount rate that justifies a predetermined conclusion. It also allows a project to be

---

[4]http://yosemite.epa.gov/ee/epa/eed.nsf/webpages/Guidelines.html/$file/Guidelines.pdf

[5]Annual rates can be found at http://www.whitehouse.gov/omb/. 2010 rates can be found at http://www.whitehouse.gov/omb/circulars_a094/a94_appx-c.

EXAMPLE 3.4

The Importance of the Discount Rate

Let's begin with an historical example. For years the United States and Canada had been discussing the possibility of constructing a tidal power project in the Passamaquoddy Bay between Maine and New Brunswick. This project would have heavy initial capital costs, but low operating costs that presumably would hold for a long time into the future. As part of their analysis of the situation, a complete inventory of costs and benefits was completed in 1959.

Using the same benefit and cost figures, Canada concluded that the project should not be built, while the United States concluded that it should. Because these conclusions were based on the same benefit–cost data, the differences can be attributed solely to the use of different discount rates. The United States used 2.5 percent while Canada used 4.125 percent. The higher discount rate makes the initial cost weigh much more heavily in the calculation, leading to the Canadian conclusion that the project would yield a negative net benefit. Since the lower discount rate weighs the lower future operating costs relatively more heavily, Americans saw the net benefit as positive.

In a more recent illustration of why the magnitude of the discount rate matters, on October 30, 2006 economist Nicholas Stern from the London School of Economics issued a report using a discount rate of 0.1 percent that concluded that the benefits of strong, early action on climate change would considerably outweigh the costs. Other economists, such as William Nordhaus of Yale University, who prefer a discount rate around 6 percent, believe that optimal economic policies to slow climate change involve only modest rates of emissions reductions in the near term, followed by sharp reductions in the medium and long term.

In this debate the desirability of strong current action is dependent (at least in part) on the size of the discount rate used in the analysis. Higher discount rates reduce the present value of future benefits from current investments in abatement, implying a smaller marginal benefit. Since the costs associated with those investments are not affected nearly as much by the choice of discount rate (remember that costs occurring in the near future are discounted less), a lower present value of marginal benefit translates into a lower optimal investment in abatement.

Far from being an esoteric subject, the choice of the discount rate is fundamentally important in defining the role of the public sector, the types of projects undertaken, and the allocation of resources across generations.

Sources: Edith Stokey and Richard Zeckhauser. *A Primer for Policy Analysis* (New York: W. W. Norton, 1978): 164–165; Raymond Mikesell. *The Rate of Discount for Evaluating Public Projects* (Washington, DC: The American Enterprise Institute for Public Policy Research, 1977): 3–5; the Stern Report: http://webarchive.nationalarchives.gov.uk/ and /http://www.hm-treasury.gov.uk/sternreview_index.htm; William Nordhaus. "A Review of the Stern Review on the Economics of Climate Change," *Journal of Economic Literature* Vol. XLV (September 2007): 686–702 considered independently of fluctuations in the true social cost of capital due to cycles in the behavior of the economy. On the other hand, when the social opportunity cost of capital differs from this administratively determined level, the benefit–cost analysis will not, in general, define the efficient allocation.

Divergence of Social and Private Discount Rates

Earlier we concluded that producers, in their attempt to maximize producer surplus, also maximize the present value of net benefits under the "right" conditions, such as the absence of externalities, the presence of properly defined property rights, and the presence of competitive markets within which the property rights can be exchanged.

Now let's consider one more condition. If resources are to be allocated efficiently, firms must use the same rate to discount future net benefits as is appropriate for society at large. If firms were to use a higher rate, they would extract and sell resources faster than would be efficient. Conversely, if firms were to use a lower-than-appropriate discount rate, they would be excessively conservative.

Why might private and social rates differ? The social discount rate is equal to the social opportunity cost of capital. This cost of capital can be separated into two components: the risk-free cost of capital and the risk premium. The *risk-free cost of capital* is the rate of return earned when there is absolutely no risk of earning more or less than the expected return. The *risk premium* is an additional cost of capital required to compensate the owners of this capital when the expected and actual returns may differ. Therefore, because of the risk premium, the cost of capital is higher in risky industries than in no-risk industries.

One difference between private and social discount rates may stem from a difference in social and private risk premiums. If the risk of certain private decisions is different from the risks faced by society as a whole, then the social and private risk premiums may differ. One obvious example is the risk *caused* by the government. If the firm is afraid its assets will be taken over by the government, it may choose a higher discount rate to make its profits before nationalization occurs. From the point of view of society—as represented by government—this is not a risk and, therefore, a lower discount rate is appropriate. When private rates exceed social rates, current production is higher than is desirable to maximize the net benefits to society. Both energy production and forestry have been subject to this source of inefficiency.

Another divergence in discount rates may stem from different underlying rates of time preference. Such a divergence in time preferences can cause not only a divergence between private and social discount rates (as when firms have a higher rate of time preference than the public sector), but even between otherwise similar analyses conducted in two different countries.

Time preferences would expected to be higher, for example, in a cash-poor, developing country than in an industrialized country. Since the two benefit–cost analyses in these two countries would be based upon two different discount rates, they might come to quite different conclusions. What is right for the developing country may not be right for the industrialized country and vice versa.

Although private and social discount rates do not always diverge, they may. When those circumstances arise, market decisions are not efficient.

A Critical Appraisal

We have seen that it is sometimes, but not always, difficult to estimate benefits and costs. When this estimation is difficult or unreliable, it limits the value of a benefit–cost analysis. This problem would be particularly disturbing if biases tended to increase or decrease net benefits systematically. Do such biases exist?

In the early 1970s, Robert Haveman (1972) conducted a major study that shed some light on this question. Focusing on Army Corps of Engineers water projects, such as flood control, navigation, and hydroelectric power generation, Haveman compared the *ex ante* (before the fact) estimate of benefits and costs with their *ex post* (after the fact) counterparts. Thus, he was able to address the issues of accuracy and bias. He concluded that

> *In the empirical case studies presented, ex post estimates often showed little relationship to their ex ante counterparts. On the basis of the few cases and the a priori analysis presented here, one could conclude that there is a serious bias incorporated into agency ex ante evaluation procedures, resulting in persistent overstatement of expected benefits. Similarly in the analysis of project construction costs, enormous variance was found among projects in the relationship between estimated and realized costs. Although no persistent bias in estimation was apparent, nearly 50 percent of the projects displayed realized costs that deviated by more than plus or minus 20 percent from ex ante projected costs.*[6]

In the cases examined by Haveman, at least, the notion that benefit–cost analysis is purely a scientific exercise was clearly not consistent with the evidence; the biases of the analysts were merely translated into numbers.

Does their analysis mean that benefit–cost analysis is fatally flawed? Absolutely not! It does, however, highlight the importance of calculating an accurate value and of including all of the potential benefits and costs (e.g., nonmarket values). It also serves to remind us, however, that benefit–cost analysis is not a stand-alone technique. It should be used in conjunction with other available information. Economic analysis including benefit–cost analysis can provide useful information, but it should not be the only determinant for all decisions.

Another shortcoming of benefit–cost analysis is that it does not really address the question of who reaps the benefits and who pays the cost. It is quite possible for a particular course of action to yield high net benefits, but to have the benefits

[6]A more recent assessment of costs (Harrington et al., 1999) found evidence of both overestimation and underestimation, although overestimation was more common. The authors attributed the overestimation mainly to a failure to anticipate technical innovation.

borne by one societal group and the costs borne by another. This admittedly extreme case does serve to illustrate a basic principle—ensuring that a particular policy is efficient provides an important, but not always the sole, basis for public policy. Other aspects, such as who reaps the benefit or bears the burden, are also important.

In summary, on the positive side, benefit–cost analysis is frequently a very useful part of the policy process. Even when the underlying data are not strictly reliable, the outcomes may not be sensitive to that unreliability. In other circumstances, the data may be reliable enough to give indications of the consequences of broad policy directions, even when they are not reliable enough to fine-tune those policies. Benefit–cost analysis, when done correctly, can provide a useful complement to the other influences on the political process by clarifying what choices yield the highest net benefits to society.

On the negative side, benefit–cost analysis has been attacked as seeming to promise more than can actually be delivered, particularly in the absence of solid benefit information. This concern has triggered two responses. First, regulatory processes have been developed that can be implemented with very little information and yet have desirable economic properties. The recent reforms in air pollution control, which we cover in Chapter 15, provide one powerful example.

The second approach involves techniques that supply useful information to the policy process without relying on controversial techniques to monetize environmental services that are difficult to value. The rest of this chapter deals with the two most prominent of these—cost-effectiveness analysis and impact analysis.

Even when benefits are difficult or impossible to quantify, economic analysis has much to offer. Policy-makers should know, for example, how much various policy actions will cost and what their impacts on society will be, even if the efficient policy choice cannot be identified with any certainty. Cost-effectiveness analysis and impact analysis both respond to this need, albeit in different ways.

Cost-Effectiveness Analysis

What can be done to guide policy when the requisite valuation for benefit–cost analysis is either unavailable or not sufficiently reliable? Without a good measure of benefits, making an efficient choice is no longer possible.

In such cases, frequently it is possible, however, to set a policy target on some basis other than a strict comparison of benefits and costs. One example is pollution control. What level of pollution should be established as the maximum acceptable level? In many countries, studies of the effects of a particular pollutant on human health have been used as the basis for establishing that pollutant's maximum acceptable concentration. Researchers attempt to find a threshold level below which no damage seems to occur. That threshold is then further lowered to provide a margin of safety and that becomes the pollution target.

Approaches could also be based upon expert opinion. Ecologists, for example, could be enlisted to define the critical numbers of certain species or the specific critical wetlands resources that should be preserved.

Once the policy target is specified, however, economic analysis can have a great deal to say about the cost consequences of choosing a means of achieving that objective. The cost consequences are important not only because eliminating wasteful expenditures is an appropriate goal in its own right, but also to assure that they do not trigger a political backlash.

Typically, several means of achieving the specified objective are available; some will be relatively inexpensive, while others turn out to be very expensive. The problems are frequently complicated enough that identifying the cheapest means of achieving an objective cannot be accomplished without a rather detailed analysis of the choices.

Cost-effectiveness analysis frequently involves an *optimization procedure*. An optimization procedure, in this context, is merely a systematic method for finding the lowest-cost means of accomplishing the objective. This procedure does not, in general, produce an efficient allocation because the predetermined objective may not be efficient. All efficient policies are cost-effective, but not all cost-effective policies are efficient.

In Chapter 2 we introduced the efficiency equimarginal principle. According to that principle, net benefits are maximized when the marginal benefit is equal to the marginal cost.

A similar, and equally important equimarginal principle exists for cost-effectiveness:

Second Equimarginal Principle (the Cost-Effectiveness Equimarginal Principle): The least-cost means of achieving an environmental target will have been achieved when the marginal costs of all possible means of achievement are equal.

Suppose we want to achieve a specific emissions reduction across a region, and several possible techniques exist for reducing emissions. How much of the control responsibility should each technique bear? The cost-effectiveness equimarginal principle suggests that the techniques should be used such that the desired reduction is achieved and the cost of achieving the last unit of emissions reduction (in other words, the marginal control cost) should be the same for all sources.

To demonstrate why this principle is valid, suppose that we have an allocation of control responsibility where marginal control costs are much higher for one set of techniques than for another. This cannot be the least-cost allocation since we could lower cost while retaining the same amount of emissions reduction. Costs could be lowered by allocating more control to the lower marginal cost sources and less to the high marginal cost sources. Since it is possible to find a way to lower cost, then clearly the initial allocation could not have minimized cost. Once marginal costs are equalized, it becomes impossible to find any lower-cost way of achieving the same degree of emissions reduction; therefore, that allocation must be the allocation that minimizes costs.

In our pollution control example, cost-effectiveness can be used to find the least-cost means of meeting a particular standard and its associated cost. Using this cost as a benchmark case, we can estimate how much costs could be expected to increase

EXAMPLE 3.5

NO_2 Control in Chicago: An Example of Cost-Effectiveness Analysis

In order to compare compliance costs of meeting a predetermined ambient air quality standard in Chicago, Seskin, Anderson, and Reid (1983) gathered information on the cost of control for each of 797 stationary sources of nitrogen oxide emissions in the city of Chicago, along with measured air quality at 100 different locations within the city. The relationship between ambient air quality at those receptors and emissions from the 797 sources was then modeled using mathematical equations. Once these equations were estimated, the model was calibrated to ensure that it was capable of re-creating the actual situation in Chicago. Following successful calibration, this model was used to simulate what would happen if EPA were to take various regulatory actions.

The results indicated that a cost-effective strategy would cost less than one-tenth as much as the traditional approach to control and less than one-seventh as much as a more sophisticated version of the traditional approach. In absolute terms, moving to a more cost-effective policy was estimated to save more than $100 million annually in the Chicago area alone. In Chapters 15 and 16 we shall examine in detail the current movement toward cost-effective polices, a movement triggered in part by studies such as this one.

from this minimum level if policies that are not cost-effective are implemented. Cost-effectiveness analysis can also be used to determine how much compliance costs can be expected to change if the EPA chooses a more stringent or less stringent standard. The case study presented in Example 3.5 not only illustrates the use of cost-effectiveness analysis, but also shows that costs can be very sensitive to the regulatory approach chosen by the EPA.

Impact Analysis

What can be done when the information needed to perform a benefit–cost analysis or a cost-effectiveness analysis is not available? The analytical technique designed to deal with this problem is called *impact analysis*. An impact analysis, regardless of whether it focuses on economic impact or environmental impact or both, attempts to quantify the consequences of various actions.

In contrast to benefit–cost analysis, a pure impact analysis makes no attempt to convert all these consequences into a one-dimensional measure, such as dollars, to ensure comparability. In contrast to cost-effectiveness analysis, impact analysis does not necessarily attempt to optimize. Impact analysis places a large amount of relatively undigested information at the disposal of the policy-maker. It is up to the policy-maker to assess the importance of the various consequences and act accordingly.

On January 1, 1970, President Nixon signed the National Environmental Policy Act of 1969. This act, among other things, directed all agencies of the federal government to

> *include in every recommendation or report on proposals for legislation and other major Federal actions significantly affecting the quality of the human environment, a detailed statement by the responsible official on—*
>
> *i. the environmental impact of the proposed action,*
>
> *ii. any adverse environmental effects which cannot be avoided should the proposal be implemented,*
>
> *iii. alternatives to the proposed action,*
>
> *iv. the relationships between local short-term uses of man's environment and the maintenance and enhancement of long-term productivity; and*
>
> *v. any irreversible and irretrievable commitments of resources which would be involved in the proposed action should it be implemented.*[7]

This was the beginning of the environmental impact statement, which is now a familiar, if controversial, part of environmental policy-making.

Current environmental impact statements are more sophisticated than their early predecessors and may contain a benefit–cost analysis or a cost-effectiveness analysis in addition to other more traditional impact measurements. Historically, however, the tendency had been to issue huge environmental impact statements that are virtually impossible to comprehend in their entirety.

In response, the Council on Environmental Quality, which, by law, administers the environmental impact statement process, has set content standards that are now resulting in shorter, more concise statements. To the extent that they merely quantify consequences, statements can avoid the problem of "hidden value judgments" that sometimes plague benefit–cost analysis, but they do so only by bombarding the policy-makers with masses of noncomparable information.

Summary

Finding a balance in the relationship between humanity and the environment requires many choices. Some basis for making rational choices is absolutely necessary. If not made by design, decisions will be made by default.

Normative economics uses benefit–cost analysis for judging the desirability of the level and composition of provided services. Cost-effectiveness analysis and impact analysis offer alternatives to benefit–cost analysis. All of these techniques offer valuable information for decision making and all have shortcomings.

[7]83 Stat. 853.

A static efficient allocation is one that maximizes the net benefit over all possible uses of those resources. The dynamic efficiency criterion, which is appropriate when time is an important consideration, is satisfied when the outcome maximizes the present value of net benefits from all possible uses of the resources. Later chapters examine the degree to which our social institutions yield allocations that conform to these criteria.

Because benefit–cost analysis is both very powerful and very controversial, in 1996 a group of economists of quite different political persuasions got together to attempt to reach some consensus on its proper role in environmental decision making. Their conclusion is worth reproducing in its entirety:

> *Benefit-cost analysis can play an important role in legislative and regulatory policy debates on protecting and improving health, safety, and the natural environment. Although formal benefit-cost analysis should not be viewed as either necessary or sufficient for designing sensible 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. If properly done, benefit-cost analysis can be of great help to agencies participating in the development of environmental, health and safety regulations, and it can likewise be useful in evaluating agency decision-making and in shaping statutes.*[8]

Even when benefits are difficult to calculate, however, economic analysis in the form of cost-effectiveness can be valuable. This technique can establish the least expensive ways to accomplish predetermined policy goals and to assess the extra costs involved when policies other than the least-cost policy are chosen. What it cannot do is answer the question of whether those predetermined policy goals are efficient.

At the other end of the spectrum is impact analysis, which merely identifies and quantifies the impacts of particular policies without any pretense of optimality or even comparability of the information generated. Impact analysis does not guarantee an efficient outcome.

All three of the techniques discussed in this chapter are useful, but none of them can stake a claim as being universally the "best" approach. The nature of the information that is available and its reliability make a difference.

Discussion Questions

1. Is risk-neutrality an appropriate assumption for benefit–cost analysis? Why or why not? Does it seem more appropriate for some environmental problems than others? If so, which ones? If you were evaluating the desirability of locating a hazardous waste incinerator in a particular town, would the Arrow-Lind rationale for risk-neutrality be appropriate? Why or why not?

[8]From Kenneth Arrow et al. "Is There a Role for Benefit-Cost Analysis in Environmental, Health and Safety Regulation?" *Science* Vol. 272 (April 12, 1996): 221–222. Reprinted with Permission from AAAS.

2. Was the executive order issued by President Bush mandating a heavier use of benefit–cost analysis in regulatory rule making a step toward establishing a more rational regulatory structure, or was it a subversion of the environmental policy process? Why?

Self-Test Exercises

1. Suppose a proposed public policy could result in three possible outcomes: (1) present value of net benefits of $4,000,000, (2) present value of net benefits of $1,000,000, or (3) present value of net benefits of –$10,000,000 (i.e., a loss). Suppose society is risk-neutral and the probability of occurrence of each of these three outcomes are, respectively, 0.85, 0.10, and 0.05, should this policy be pursued or trashed? Why?
2. a. Suppose you want to remove ten fish of an exotic species that have illegally been introduced to a lake. You have three possible removal methods. Assume that q_1, q_2, and q_3 are, respectively, the amount of fish removed by each method that you choose to use so that the goal will be accomplished by any combination of methods such that $q_1 + q_2 + q_3 = 10$. If the marginal costs of each removal method are, respectively, $10q_1$, $5q_2$, and $2.5q_3$, how much of each method should you use to achieve the removal cost-effectively?
   b. Why isn't an exclusive use of method 3 cost-effective?
   c. Suppose that the three marginal costs were constant (not increasing as in the previous case) such that MC_1=$10, MC_2=$5, and MC_3=$2.5. What is the cost-effective outcome in that case?
3. Consider the role of discount rates in problems involving long time horizons such as climate change. Suppose that a particular emissions abatement strategy would result in a $500 billion reduction in damages 50 years into the future. How would the maximum amount spent now to eliminate those damages change if the discount rate is 2 percent, rather than 10 percent?

Further Reading

Freeman, A. Myrick III. *The Measurement of Environmental and Resource Values*, 2nd ed. (Washington, DC: Resources for the Future, Inc., 2003). A comprehensive and analytically rigorous survey of the concepts and methods for environmental valuation.

Hanley, Nick, and Clive L. Spash. *Cost-Benefit Analysis and the Environment* (Brookfield, VT: Edward Elgar Publishing Company, 1994). An up-to-date account of the theory and practice of this form of analysis applied to environmental problems. Contains numerous specific case studies.

Norton, Bryan, and Ben A. Minteer. "From Environmental Ethics to Environmental Public Philosophy: Ethicists and Economists: 1973–Future," in T. Tietenberg and H. Folmer, eds. *The International Yearbook of Environmental and Resource Economics: 2002/2003* (Cheltenham, UK: Edward Elgar, 2002): 373–407. A review of the interaction between environmental ethics and economic valuation.

Scheraga, Joel D., and Frances G. Sussman. "Discounting and Environmental Management," in T. Tietenberg and H. Folmer, eds. *The International Yearbook of Environmental and Resource Economics 1998–1999* (Cheltenham, UK: Edward Elgar, 1998): 1–32. A summary of the "state of the art" for the use of discounting in environmental management.

Additional References and Historically Significant References are available on this book's Companion Website: http://www.pearsonhighered.com/tietenberg/

Appendix

The Simple Mathematics of Dynamic Efficiency*

Assume that the demand curve for a depletable resource is linear and stable over time. Thus, the inverse demand curve in year t can be written as

$$P_t = a - bq_t \tag{1}$$

The total benefits from extracting an amount q_t in year t are then the integral of this function (the area under the inverse demand curve):

$$(\text{Total benefits})_t = \int_0^{qt} (a - bq)dq \tag{2}$$

$$= aq_t - \frac{b}{2}q_t^2$$

Further assume that the marginal cost of extracting that resource is a constant c and therefore the total cost of extracting any amount q_t in year t can be given by

$$(\text{Total cost})_t = cq_t \tag{3}$$

If the total available amount of this resource is $\bar{Q}$, then the dynamic allocation of a resource over n years is the one that satisfies the maximization problem:

$$\text{Max}_q \sum_{i=1}^{n} \frac{aq_i - bq_i^2/2 - cq_i}{(1 + r)^{i-1}} + \lambda \left[\bar{Q} - \sum_{i=1}^{n} q_i\right] \tag{4}$$

Assuming that $\bar{Q}$ is less than would normally be demanded, the dynamic efficient allocation must satisfy

$$\frac{a - bq_i - c}{(1 + r)^{i-1}} - \lambda = 0, \quad i = 1, \ldots, n \tag{5}$$

$$\bar{Q} - \sum_{i=1}^{n} q_i = 0 \tag{6}$$

An implication of Equation 5 is that ($P - MC$) increases over time at rate r. This difference, which is known as the marginal user cost, will play a key role in our thinking about allocating depletable resources over time. A fuller understanding of this concept is provided in Chapter 5.

*Greater detail on the mathematics of constrained optimization can be found in any standard mathematical economics text.