Cost–Benefit Analysis
Assessing Efficiency

Cost–benefit analysis (CBA), a technique for systematically estimating the efficiency impacts of policies, came into common use in the evaluation of flood-control projects in the 1930s. It has since been mandated and applied, with varying degrees of success, across a broad spectrum of public policies. Both President Reagan’s Executive Order 12291 and President Clinton’s Executive Order 12866 required U.S. federal agencies to prepare regulatory impact analyses for any regulations that are likely to result in significant economic impact. The Clinton executive order required that analyses identify social costs and benefits and attempt to determine whether the proposed benefits of the regulation “justify” the costs to society. Congress has also sometimes required cost–benefit-like analysis in a variety of legislation, such as the Unfunded Mandates Reform Act of 1995. The federal courts in the United States now utilize a form of “cost–benefit balancing” authorized by legislation. While the evidence suggests that many federal agencies have had difficulty actually implementing CBA in practice, the requirement for such

1The U.S. Flood Control Act of 1936 required that water resource projects be evaluated on the difference between estimated benefits and costs. The Bureau of the Budget set out its own criteria in 1952, in Budget Circular A-47. For an overview, see Peter O. Steiner, “Public Expenditure Budgeting,” in Alan S. Blinder et al., The Economics of Public Finance (Washington, DC: Brookings Institution, 1974), 241–357.

Part IV: Doing Policy Analysis

analysis is now well entrenched.\(^3\) Many states now also require that regulatory initiatives include some form of CBA.\(^4\)

The wider use of CBA has become possible because economists have developed techniques for monetizing impacts occurring outside markets. For example, environmental economists now commonly use stated preference methods, such as the contingent valuation survey, to provide a basis for valuing public goods and externalities not traded in markets, and economists in a number of other fields use hedonic pricing models to assess people's willingness to pay for reductions in mortality risk. These developments have facilitated the application of CBA beyond infrastructure investments and environmental policies to social policies.\(^5\)

The appropriateness of CBA as a decision rule depends on whether efficiency is the only relevant value and the extent to which important impacts can be monetized. When values other than efficiency are relevant, CBA can still be useful as a component of multigoal policy analysis, which we discuss more fully in Chapter 15. When important impacts cannot be reasonably monetized, the first step of CBA—identifying impacts and categorizing them as costs or benefits—can be embedded in a broader analytical approach. Thus, the value to analysts of familiarity with the basic elements of CBA goes beyond its direct use as a decision rule.

In this chapter we introduce the basic elements of CBA. A thorough treatment of all the relevant theoretical and practical issues requires an entire text.\(^6\) Here we focus only on the key concepts for doing and assessing CBAs.

A Preview: Increasing Alcohol Taxes

Would a higher excise tax on alcohol be efficient? That is, would it increase aggregate social welfare? This is the sort of question that we try to answer with CBA analysis. We compare an increased tax to the current, or status quo, tax regime. We begin by identifying all the impacts of the tax. Direct impacts include higher prices of alcohol for consumers and increased revenue for the government. But the higher prices also lead to a reduction in the consumption of alcohol. Lower levels of consumption produce several indirect effects: fewer fatalities and injuries and less property damage from alcohol-related automobile accidents, and perhaps better health and increased productivity for people who reduce their consumption of alcohol.

The next task in doing a CBA of an increased tax on alcohol is to put dollar values on its impacts. For example, how much money would we have to give consumers to make them willingly accept the higher price of alcohol? (The answer is one of the costs of the tax increase.) How much would people be willing to pay to reduce their risks of being in automobile accidents? (The answer is one of the benefits of the tax increase.) After answering questions such as these for all of the impacts, we then

---


compare the marginal benefits of the tax increase with its marginal costs. If benefits exceed costs, then we conclude that the tax increase is efficient—that is, it would at least be possible to compensate fully all those who bear net costs and still have some excess left to make some people better off.

Later in this chapter we present an actual CBA of an increased tax on alcohol. First, however, we set out the basic concepts that help correctly identify, measure, and compare costs and benefits. The following sections discuss these concepts in terms of four basic steps: (1) identifying relevant impacts, (2) monetizing impacts, (3) discounting for time and risk, and (4) choosing among policies. We organize our discussion of the basic concepts underlying CBA according to these steps.

### Identifying Relevant Impacts

The first step in CBA is to identify all the impacts of the policy under consideration and categorize them as either costs or benefits for various groups. This immediately raises a key issue: Who has standing? That is, whose utility should we count when assessing costs and benefits? The question almost always arises in the context of choosing geographic boundaries. It also arises, however, when persons either cannot articulate their preferences or articulate preferences that society considers invalid.

#### Geographic Extent

The most inclusive definition of society encompasses all people, no matter where they live or to which government they owe allegiance. Analysts working for the United Nations or some other international organization might appropriately adopt such a universalistic perspective. Analysts employed by their national governments, however, would most likely view their fellow citizens, perhaps including residents of their countries who are not citizens, as the relevant societies for considering economic efficiency. Impacts that accrue outside national boundaries are typically ignored for purposes of measuring changes in economic efficiency. Of course, these external impacts may have political implications, such as when a dam in one country reduces the flow of water in a river entering another country. Because the political importance of externalities cannot always be readily determined, it is usually best to begin by listing all identifiable impact categories, whether they are internal or external to the national society. Explicit judgments can then be made about which externalities should be ignored, which monetized (usually a heroic task!), and which highlighted as “other considerations.”

The issue of geographic standing often arises for analysts working for subnational governments. For example, consider a city that is deciding whether to build a convention center. Assume that a CBA from the social perspective (giving standing to

---

7Our treatment of this topic benefited greatly from Dale Whittington and Duncan MacRae, Jr., “The Issue of Standing in Cost–benefit Analysis,” *Journal of Policy Analysis and Management* 5(4) 1986, 663–82. The term standing has its origins in the legal doctrine “standing to sue”: plaintiffs have standing if they have a legally protectable and tangible interest at stake in the litigation. See *Black's Law Dictionary*, 5th ed. (St. Paul, MN: West, 1979), at 1260.

8Some would argue that even future generations should be included in the definition of society. We will address the implications of this view in our discussion of discounting.
everyone in the country) predicts that the project will generate costs in excess of benefits of $1 million. Also assume that the national government will pay $2 million of the costs of this particular project through an intergovernmental grants program. Because the residents of the city contribute a negligible fraction of the total taxes collected by the national government, the grant appears to city residents as a $2 million benefit offsetting $2 million in costs. Thus, from the perspective of the city, the convention center appears to generate $1 million in net benefits rather than $1 million in net costs. The fact that national governments typically contribute large grants is one reason that cities compete so enthusiastically for the right to host the Olympic Games.

A variety of externalities, or spillovers, can cause a divergence between aggregate social welfare and the welfare of local governments. We can divide them into two broad categories: fiscal and economic. Fiscal externalities, like the intergovernmental grant in the example, transfer wealth, or rents, across the boundaries of local jurisdictions. In a CBA from the perspective of aggregate national social welfare, a fiscal externality appears as offsetting costs and benefits to different groups in society and, therefore, does not influence the final calculation of net benefits. For example, a program that lures a firm from one city to another typically involves costs for one city that offset the gross benefits accruing to the other.

Economic externalities, as discussed in Chapter 5, directly affect the capability of some persons outside the jurisdiction to either produce or consume. For example, if an upstream city improves its sewage treatment plant, then cities downstream enjoy cleaner water, which may enhance the recreational value of the river and perhaps reduce their cost of producing potable water. Although the upstream city might not count these downstream benefits in its CBA of the improved treatment plant, they should be included in a CBA from the perspective of social welfare.

How should analysts treat costs and benefits that are external to their clients’ jurisdictions? We believe that analysts should estimate costs and benefits from the perspectives of their clients and society. The CBA from the perspective of society indicates what should be done in an ideal world. The CBA from the perspective of the local jurisdiction indicates what should be done to serve the direct interests of the client’s constituency and, therefore, the political interests of the client. As an analyst, if you do not note significant externalities, you are failing in your duty to inform your client about appropriate values; if you do not clearly indicate the costs and benefits accruing to your client’s constituency, you are probably failing in your responsibility to represent your client’s interests.

Persons and Preferences

Should all persons within a jurisdiction have standing? Almost everyone would agree that, beyond citizens, costs and benefits incurred by legal residents should be counted. What about those incurred by illegal aliens? The answer may depend on the nature of the costs and benefits. For example, if they result from direct changes in health and safety, we might be more inclined to count them than if they result from changes in incomes. The same reasoning probably applies to citizens convicted of serious crimes. Obviously, however, such questions about standing raise difficult ethical issues. Indeed, when such questions are central to the identification of relevant impacts, multigoal analysis, rather than CBA, is likely to be the more appropriate method.

The issue of standing also arises with respect to the expression of preferences. Do families and other institutions adequately articulate what would be the informed
preferences of children, those who are mentally incompetent, or other persons with limited ability to reason? Do markets and other institutions adequately express the preferences of future generations for such amenities as environmental quality? In other words, should we give special attention to impacts on any particular groups within society? Also, should we accept all preferences as valid? For example, burglars would undoubtedly view reductions in the monetary take from their crimes as a cost of an enforcement policy. Yet, most analysts would not include this cost in their evaluation of the policy. As with the other questions about standing, when these issues become central, we should carefully reconsider the appropriateness of CBA as the method of evaluation.

Reprise of Standing Issues

The first step in CBA is to identify all relevant impacts and classify them as either costs or benefits for various groups. It is best to start out by being inclusive of any affected groups. Reasoned arguments can then be made to exclude those groups that you believe should not have standing. When a client has a subnational constituency, it is usually desirable to estimate costs and benefits from the perspectives of both the national society and the constituency. If it appears that issues of standing will be central to your analysis, then you should consider switching from the CBA to multigoal analysis.

Monetizing Impacts

The basic principle underlying CBA is the Kaldor-Hicks criterion: a policy should be adopted only if those who will gain could fully compensate losers and still be better off. In other words, when efficiency is the only relevant value, a necessary condition for adopting a policy is that it has the potential to be Pareto improving. As we discussed in Chapter 4, policies that increase social surplus are potentially Pareto improving and, therefore, meet the Kaldor-Hicks criterion. Further, when considering mutually exclusive policies, the one that produces the greatest increase in social surplus should be selected because, if adopted, it would be possible through the payment of compensation to make everyone at least as well off as they would have been under any of the alternative policies.

Many economists treat the move from the Pareto-improving criterion to the potentially Pareto-improving criterion as if it were a minor step. It is not. Actual Pareto-improving exchanges are voluntary and, by definition, make someone better off without making anyone else worse off. Potential Pareto-improving moves do not guarantee that no one is made worse off—just that everyone could be made better off
Part IV: Doing Policy Analysis

with appropriate redistribution following the move. Thus, while the Kaldor-Hicks criterion is ingenious in that it is “Pareto-like,” it is also controversial. As Richard Posner, among others, forcefully argues, CBA based on the Kaldor-Hicks criterion contains elements of utilitarianism (“the surplus of pleasure over pain-aggregated across all of the inhabitants of society”) that the pure Pareto criterion avoids.\footnote{Richard Posner, The Economics of Justice (Cambridge, MA: Harvard University Press, 1983), 49. For a discussion of the same point from a very different perspective, see Steven Kelman, “Cost–Benefit Analysis: An Ethical Critique,” Regulation 1981(January/February), 33–40.} Most fundamentally, the aggregation underlying the Kaldor-Hicks criterion implicitly compares utility across individuals in terms of a common dollar measure or money metric. Nevertheless, the criterion provides a good test for determining if policies could be efficient, if not a fully acceptable rule, for choosing among policies.

In practice, two related concepts serve as guides for estimating changes in social surplus and, therefore, for applying the Kaldor-Hicks criterion: opportunity cost and willingness to pay. They provide ways of monetizing costs and benefits.

### Valuing Inputs: Opportunity Cost

Public policies usually require resources (factor inputs) that could be used to produce other goods. Some examples: public works projects such as dams, bridges, and highways require labor, materials, land, and equipment; social service programs typically require professional services and office space; and wilderness preserves, recreational areas, and parks require at least land. The resources used to implement these policies cannot be used to produce other goods. The values of the forgone goods measure the costs of policies. In general, the opportunity cost of a policy is the value of the required resources in their best alternative uses.

The nature of the market for a resource determines how we go about measuring its opportunity cost. Three situations arise: (1) the market for the resource is efficient (no market failures), and purchases for the project will have a negligible effect on price (constant marginal costs); (2) the market for the resource is efficient, but purchases for the project will have a noticeable effect on price (rising or falling marginal costs); and (3) the market for the resource is inefficient (market failure).

#### Efficient Markets and Negligible Price Effects.

In an efficient market, the equilibrium price equals the marginal social cost of production. The amount that must be paid to purchase one additional unit exactly equals the opportunity cost of that unit. Because marginal cost is constant (the supply curve is perfectly elastic), we can purchase additional units at the original price. The opportunity cost of the marginal units is simply the total amount we spend to purchase them.

Panel (a) in Figure 16.1 illustrates the opportunity cost of purchases in efficient factor markets with constant marginal cost. Purchasing $Q'$ units of the factor for a public project can be thought of as shifting the demand schedule for the factor, $D$, to the right by a horizontal distance of $Q'$. (Strictly speaking, $D$ is a derived demand schedule—it represents the marginal valuations of various quantities of the factor that derive from their use in producing goods that consumers directly demand.) Because the supply schedule $S$ is perfectly elastic, and marginal cost ($MC$) is constant, price remains at $P_0$. The total amount spent on the factor used by the project is $P_0$ times $Q'$, which equals the area of the shaded rectangle $ab(Q_0 + Q')Q_0$, the total social cost of
Chapter 16: Cost–Benefit Analysis

Figure 16.1 Measuring Opportunity Cost in Efficient Factor Markets

(a) Constant Marginal Cost: Perfectly Elastic Supply

(b) Rising Marginal Cost: Upward Sloping Supply
using $Q'$ units of the factor for the public project. If the $Q'$ units are not used for the project, then $P_0$ times $Q$ worth of goods could be produced elsewhere in the economy. Thus, this public expenditure exactly equals the opportunity cost of using $Q'$ units of the factor for the project.

Because most factors have neither steeply rising nor declining marginal cost schedules, interpreting expenditures as opportunity costs is usually reasonable when the quantity purchased makes a small addition to the total demand for the factor. For example, consider a proposed remedial reading program for a school district. The additional textbooks are purchased in a national market. They represent only a small addition to total demand for textbooks and, hence, result in a negligible increase in their price. In contrast, to hire qualified reading teachers for the program in the local labor market may require higher salaries than those being paid to the reading teachers already employed.

**Efficient Markets with Noticeable Price Effects.** Panel (b) in Figure 16.1 illustrates the effect of factor purchases when marginal costs are increasing so that the supply schedule is upward sloping. As in panel (a), purchase of $Q'$ units of the factor for use by the public project shifts the demand schedule to the right. Because the supply schedule $S$ is upward sloping, the equilibrium price rises from $P_0$ to $P_1$. The total expenditure on the $Q'$ units of the factor needed for the project is $P_1$ times $Q'$, which equals the area of rectangle $Q_2deQ_1$.

Unlike the case in which marginal costs are constant, this expenditure is not the opportunity cost of using $Q'$ units of the factor for the project. When purchases for a project induce a price change in a factor market, the effects of the price change on social surplus within the market must be taken into account when calculating opportunity cost. The general rule is that *opportunity cost equals expenditure less (plus) any increase (decrease) in social surplus occurring in the factor market itself*. In other words, expenditures do not accurately represent opportunity costs when purchases cause social surplus changes in factor markets.

Referring again to panel (b) in Figure 16.1, we can identify the changes in producer and consumer surplus that result from the increase in price from $P_0$ to $P_1$. Producer surplus increases by the area of trapezoid $P_1ecP_0$, the increase of the area below price and above the supply schedule. At the same time, consumer surplus decreases by the area of trapezoid $P_1deP_0$ (the decrease in the area above price and below the demand schedule). Subtracting the loss in consumer surplus from the gain in producer surplus leaves a net gain of social surplus in the factor market equal to the area of triangle $cde$. Subtracting this social surplus gain from the expenditure on the $Q'$ units of the factor needed for the project yields the opportunity cost represented by the shaded area within the geometric figure $Q_2decQ_1$. Note that calculation of this area is straightforward when the supply and demand schedules are linear, it is the amount of the factor purchased for the project, $Q'$, multiplied by the average of the old and new prices, $1/2(P_1 + P_0)$.

An alternative explanation may be helpful in understanding why the shaded area represents opportunity cost. Imagine that the government obtains the $Q'$ units of the output by first restricting supply to the market from $Q_0$ to $Q_2$ and then ordering

---

12Exactly one-half of the area of rectangle $degf$ is shaded. (With a bit of geometry one can show that the area of triangle $cdf$ plus the area of triangle $ceg$ equals the area of triangle $cde$.) Therefore, the total shaded area equals $1/2(P_1 - P_0)Q'$ plus $P_0Q'$, which equals $1/2(P_1 + P_0)Q'$, the average price times the quantity purchased.
the firms in the industry to produce $Q'$ for the government at cost. The social surplus loss resulting from restricting market supply to $Q_2$ is the area of triangle $cdh$, the deadweight loss. The total cost of producing the $Q'$ units for the government is the area of trapezoid $Q_2Q_1$. Adding these areas gives the same opportunity cost as calculated under the assumption that the government purchases the factor like any other participant in the market. Note, however, that while public expenditures on the factor exceed opportunity cost when the government purchases it like everyone else, expenditure falls short of opportunity cost when the government obtains the factor through directives. In other words, budgetary cost can either understate or overstate social opportunity cost.

**Inefficient Markets.** In an efficient market, price equals marginal cost. If price does not equal marginal cost, then the result is allocative inefficiency. As we saw in Chapters 5, 6, 8, and 9, a variety of circumstances can lead to inefficiency: absence of a working market, market failures (public goods, externalities, natural monopolies, and information asymmetries), markets with few sellers, and distortions due to government interventions (such as taxes, subsidies, price ceilings, and price floors). Any of these distortions can arise in factor markets, complicating the estimation of opportunity cost.

Consider a proposal to establish more courts so that more criminal trials can be held. Budgetary costs will include the salaries of judges and court attendants, rent for courtrooms and offices, and perhaps expenditures for additional correctional facilities (because the greater availability of trial capacity may permit more vigorous prosecution). The budget may also show payments to jurors. Typically, however, these payments just cover commuting expenses. Compensation paid to jurors for their time is typically not related to their wage rate but, rather, set at a nominal per diem rate. Thus, the budgetary outlay for payments to jurors almost certainly underestimates the opportunity cost of the jurors’ time. A better estimate of opportunity cost would be commuting expenses plus the number of juror-hours times either the average or the median hourly wage rate (including benefits) for the locality. The commuting expenses estimate the resource costs of transporting the jurors to the court; the hourly wage rate times the hours spent on jury duty estimates the value of goods forgone because of lost labor.

Assessing opportunity costs in the presence of market failures or government interventions requires a careful accounting of social surplus changes. For example, let us examine the opportunity costs of labor in a market where either minimum wage laws or union bargaining power keeps the wage rate above the market clearing level. In Figure 16.2, the pre-project demand schedule for labor $D$ and the supply schedule for labor $S$ intersect at $W_E$, the equilibrium wage in the absence of the wage floor, $W_M$. At the wage floor, $L_S$ units of labor are offered but only $L_D$ units are demanded so that $L_S - L_D$ units are “unemployed.” Now imagine that $L'$ units are hired for the project. This shifts the demand schedule to the right by $L'$. As long as $L'$ is less than the quantity of unemployed labor, price will remain at the floor. The total expenditure on labor for the project is $W_M$ times $L'$, which equals the area of rectangle $abL'L_D$. But trapezoid $abcd$ represents producer surplus enjoyed by the newly hired and, hence, should be subtracted from the expenditure to obtain an opportunity cost equal to the shaded area inside trapezoid $cdL'DT$. Alternatively, we can think of the shaded area as the value of the leisure time (a good) given up by the newly hired workers.

The interpretation of $cdL'DT$ as the opportunity cost of the labor depends on the assumption that the workers hired by the project place the lowest values on
leisure among all those unemployed. (They have the lowest reservation wages, the minimum wages at which they will offer labor.) Note that this may not be a reasonable assumption because at wage $W_M$ all those offering labor between $L_D$ and $L_S$ will try to get the jobs created by the project. Consequently, the shaded area in Figure 16.2 is the minimum opportunity cost; the maximum opportunity cost would result if the $L'$ workers hired were the unemployed whose reservation wages were near $W_M$. Actual opportunity cost would likely be some average of these two extremes. In any event, the opportunity cost of the project labor would be below its budgetary cost.

Other market distortions affect opportunity costs in predictable ways. In factor markets where supply is taxed, expenditures overestimate opportunity costs; in factor markets where supply is subsidized, expenditures underestimate opportunity costs. In factor markets exhibiting positive supply externalities, expenditures overestimate opportunity costs; in factor markets exhibiting negative supply externalities, expenditures underestimate opportunity costs. In monopolistic factor markets, expenditures overestimate opportunity costs. To determine opportunity costs in these cases, apply the general rule: opportunity cost equals expenditures on the factor minus (plus) gains (losses) in social surplus occurring in the factor market.

A final point on opportunity costs: the relevant determination is what must be given up today and in the future, not what has already been given up, which are sunk costs. For instance, suppose that you are asked to reevaluate a decision to build a bridge after construction has already begun. What is the opportunity cost of the steel and concrete that are already in place? It is not the original expenditure made to purchase them. Rather, it is the value of these materials in the best alternative...
use—most likely measured by the maximum amount that they could be sold for as scrap. Conceivably, the cost of scrapping the materials may exceed their value in any alternative use so that salvaging them would not be justified. Indeed, if salvage is necessary, say for environmental or other reasons, then the opportunity cost of the material will be negative (and thus counted as a benefit) when calculating the net benefits of continuing construction. In situations where resources that have already been purchased have exactly zero scrap value (the case of labor already expended, for instance), the costs are sunk and are not relevant to our decisions concerning future actions.

**Valuing Outcomes: Willingness to Pay**

The valuation of policy outcomes is based on the concept of willingness to pay: benefits are the sum of the maximum amounts that people would be willing to pay to gain outcomes that they view as desirable; costs are the sum of the maximum amounts that people would be willing to pay to avoid outcomes that they view as undesirable. Of course, estimating changes in social surpluses that occur in relevant markets enables us to take account of these costs and benefits.

**Efficient Markets.** Valuation is relatively straightforward when the policies under consideration will affect the supply schedules of goods in efficient markets. A general guideline holds for assessing benefits: the benefits of a policy equal the net revenue generated by the policy plus the social surplus changes in the markets in which the effects of the policy occur. Note that the benefits can be either positive or negative. We generally refer to those that are negative as costs. Indeed, if we think of the use of factor inputs as an impact, then this statement of benefit calculation encompasses opportunity costs. In other words, depending on how we initially categorize policy impacts, we may measure them either as a cost or as a negative benefit.

Two situations are common: First, the policy may directly affect the quantity of some good available to consumers. For example, opening a publicly operated day-care center shifts the supply schedule to the right—more day care is offered to consumers at each price. Second, the policy may shift the supply schedule by altering the price or availability of some factor used to produce the good. For example, deepening a harbor so that it can accommodate large, efficient ships reduces the costs of transporting bulk commodities to and from the port.

Figure 16.3 shows the social surplus changes that result from additions to supply. The intersection of the demand schedule $D$ and the supply schedule $S$ indicates the equilibrium price $P_0$ prior to the project. If the project directly adds a quantity $Q$ to the market, then the supply schedule, as seen by consumers, shifts from $S$ to $S + Q$, and the equilibrium price falls to $P_1$. If consumers must purchase the additional units from the project, then the gain in consumer surplus equals the area of trapezoid $P_0abP_1$. Because suppliers continue to operate on the original supply schedule, they suffer a loss of surplus equal to the area of trapezoid $P_0acP_1$ so that the net change in social surplus equals the area of triangle $abc$, which is darkly shaded. In addition, however, the project enjoys a revenue equal to $P_1$ times $Q$, the area of rectangle $Q_2CBQ_1$. So the sum of project revenues and the change in social surplus in the market equals the area of trapezoid $Q_2cabQ_1$, which is the total benefit from the project selling $Q$ units in the market.
What benefits accrue if the $Q$ units are instead distributed free to selected consumers? If the $Q$ units are given to consumers who would have purchased an identical or greater number of units at price $P_1$, then the benefit measure is exactly the same as when the project’s output is sold. Even though no project revenues accrue, consumers enjoy a surplus that is greater by the area of rectangle $Q_2cbQ_1$, which exactly offsets the revenue that would accrue if the project’s output is sold.

If the $Q$ units are distributed to consumers in greater quantities than they would have purchased at price $P_1$, then area $Q_2cabQ_1$ will be the project’s benefit only if these recipients can and do sell the excess quantities to those who would have bought them. If the recipients keep any of the excess units, then area $Q_2cabQ_1$ will overestimate the project’s benefit in two ways. First, in contrast to the situation in which the $Q$ units are sold, some consumers will value their marginal consumption at less than $P_1$. (If they valued it at or above $P_1$, they would have been willing to purchase at $P_1$.) Second, their added consumption shifts the demand schedule to the right so that the market price after provision of $Q$ units by the project will not fall all the way to $P_1$. Even if the recipients do not keep any of the excess, the project’s benefits may be smaller than the area of $Q_2cabQ_1$ because of transactions costs.

For example, suppose that the project provides previously stockpiled gasoline to low-income consumers during an oil supply disruption (an in-kind subsidy). Some of

![Figure 16.3 Measuring Benefits in an Efficient Market](image-url)
the recipients will find themselves with more gasoline than they would have purchased on their own at price $P_1$; therefore, they may try to sell the excess. Doing so will be relatively easy if access to the stockpiled gasoline is provided through legally transferable coupons; it would be much more difficult if the gasoline had to be physically taken away by the recipients. If the gasoline can be costlessly traded among consumers, then we would expect the outcome to be identical to the one that would result if the gasoline were sold in the market, with the revenue given directly to the low-income consumers.

Next suppose that the project, like the harbor deepening, lowers the cost of supplying the market. In this case, the supply schedule as seen by both consumers and producers shifts to $S + Q$, not because the project directly supplies $Q$ to the market but, rather, because reductions in the marginal costs of firms allow them to offer $Q$ additional units profitably at each price along $S + Q$. As with the case of direct supply of $Q$, the new equilibrium price is $P_1$ and consumers gain surplus equal to the area of trapezoid $P_0abP_1$. Producers gain surplus equal to the difference in the areas of triangle $P_0dP_1$ (the producer surplus with supply schedule $S$) and triangle $P_1bd$ (the producer surplus with supply schedule $S + Q$). Area $P_1ce$ is common to the two triangles and therefore cancels, leaving the area $cbde$ minus the area $P_0acP_1$. Adding this gain to the gain in consumer surplus, which can be stated as area $P_0acP_1$ plus $abc$, leaves areas $abc$ plus $cbde$. Thus, the gain in social surplus resulting from the project equals the area of trapezoid $abde$. Because no project revenue is generated, area $abde$ alone is the benefit of the project.

**Distorted Markets.** If market failures or government interventions distort the relevant product market, then the determination of the benefits and costs of policy effects is more difficult. Although the same general rule for measuring benefits continues to apply, complications arise in determining the correct social surplus changes in distorted (inefficient) markets. For example, a program that subsidizes the purchase of rodent extermination services in a poor neighborhood probably has an external effect: the fewer rodents in the neighborhood, the easier it is for residents in adjoining neighborhoods to keep their rodent populations under control. Thus, in Figure 16.4 we show the market demand schedule $DM$ as understating the social demand schedule $DS$ so that the market equilibrium price $P_0$ and quantity $Q_0$ are too low from the social perspective.

What are the social benefits of a program that makes vouchers worth $v$ dollars per unit of extermination service available to residents of the poor neighborhood? When the vouchers become available, residents of the poor neighborhood see a supply schedule that is lower than the market supply schedule by $v$ dollars. Consequently, they increase their purchases of extermination services from $Q_0$ to $Q_1$, and they see an effective price equal to the new market price less the per unit subsidy ($P_1 - v$). Consumers in the targeted neighborhood enjoy a surplus gain equal to the area of trapezoid $P_0d(P_1 - v)$; producers, who now see a price of $P_1$, receive a surplus gain equal to the area of trapezoid $P_1edP_0$, and people in the surrounding neighborhoods, who enjoy the positive externality, gain surplus equal to the area of figure $abcd$, the area between the market and social demand schedules over the increase in consumption. Ignoring administrative costs, the program must pay out $v$ times $Q_1$, which equals the area of rectangle $P_1ec(P_1 - v)$. Subtracting this program cost from the gain in social surplus in the market yields net program benefits equal to the area of trapezoid $abde$. 

Chapter 16: Cost–Benefit Analysis 395
Estimating the Demand for Nonmarketed Goods

Public projects often produce goods that are not traded in markets. Indeed, market failures involving public goods provide the primary rationale for direct public supply. Consequently, we frequently encounter situations in which we cannot directly infer demand schedules from market data. There are a number of general approaches for assessing the demand for goods that are not traded in markets. Here we briefly sketch three of these approaches: hedonic price models, survey assessments, and activity surveys.


See Boardman et al., Cost–Benefit Analysis, Chapters 11–14, for explanations of general approaches to benefit estimation.
Hedonic Price Models. Levels of nonmarketed goods sometimes affect the prices of goods that are traded in markets. For example, a housing price not only reflects the characteristics of the house but also such locational factors as the quality of the public school district, the level of public safety, and accessibility. (Remember the old saying: "There are three important factors in real estate—location, location, and location.") Now, if we could find houses that were identical in all these factors except, say, the level of public safety, then we could interpret any price difference as the value the market places on the difference in safety levels.

In practice, we are almost never fortunate enough to find goods that are identical in every way except for the particular characteristic of interest. Nevertheless, statistical techniques can often be used to identify the independent contribution of specific characteristics on price. The theoretical foundation for such estimation is known as the hedonic price model.\textsuperscript{15} Later in this chapter we discuss use of the hedonic price model to estimate the value of life implicit in risk–wage trade-offs. Other applications include using intercity salary differences to estimate the benefits of air quality improvements,\textsuperscript{16} using housing values to estimate the value of air quality improvements,\textsuperscript{17} housing price differences to estimate implicit values placed on health risks,\textsuperscript{18} and the quality of public schools.\textsuperscript{19}

Lack of appropriate data severely limits the widespread applicability of the hedonic price model. Unless data describing all the major characteristics affecting price are available, a reliable estimate of the independent contribution of the characteristic of interest cannot be made. Even when data on all major characteristics are available, it may still be difficult to separate the independent effects of characteristics that tend to vary in the same pattern. Nonetheless, the hedonic price model offers a conceptually attractive approach even if its practical applicability is limited.

Stated Preferences (Contingent Valuation Surveys). A direct approach to estimating the benefits of public goods is to ask a sample of people how much they would be willing to pay to obtain them in \textit{contingent valuation surveys}.\textsuperscript{20} By comparing the demographic characteristics of the sample to those of the general population, an estimate of the aggregate willingness to pay for specific levels of public goods can be made. A major advantage of this approach is that it permits estimation of the benefits of a wide range of public goods, including national ones like defense that do not vary at the local level. It also permits estimates of the benefits people derive from the provision of public goods to others. In some contexts, where there are no direct or easily


observable indirect behavioral traces, contingent valuation surveys may be the only feasible way to estimate willingness to pay.21

Obviously, this approach suffers from all the well-known problems of survey research: answers are sensitive to the particular wording of questions; nonrandom sampling designs or nonresponses can lead to unrepresentative samples; respondents have limited attention spans; and respondents often have difficulty putting hypothetical questions into meaningful contexts. There are also problems specific to valuing public goods. First, there is the practical difficulty of explaining the good being valued adequately to respondents. If respondents do not understand what is being valued, then it is not clear how to interpret estimates of their willingness to pay. Second, respondents may not treat the hypothetical choice they face as an economic decision. If this is the case, then they may ignore budget constraints and respond to gain moral satisfaction, or what has been more colloquially called a warm glow feeling.22 Third, respondents may answer strategically. For example, because I know that I will not actually have to pay the amount that I state, I may inflate my true willingness to pay for public goods I prefer, so as to increase the chances that the survey will value them highly. Although a conceptual threat to surveys, strategic behavior is probably not that serious in practice.23 Further, strategic behavior can be reduced by using referendum-type formats for valuation—typically with each respondent being given a random dollar amount and asked if he or she would vote for provision of the public good if it were to cost that amount.

Despite concerns about relying on stated rather than revealed preferences, contingent valuation is now widely used. A blue-ribbon panel of economists has endorsed its limited use in environmental damage estimation.24 It is also accepted as evidence in federal court cases.25

Activity Surveys: The Travel Cost Method. Some of the problems associated with opinion surveys can be avoided by surveying people about their actual behavior. The most common application of this approach is the estimation of the value of recreation sites from people’s use patterns.26 For example, imagine that we wish to estimate the value people place on a regional park. We could survey people who live various distances from the park about how frequently they visit it. We then would statistically relate the frequency of use to the travel costs (the effective price of using the park) and the demographic characteristics of the respondents. These relationships enable us to estimate the population’s demand schedule for park visits so that we can apply standard consumer surplus analysis. Of course, the accuracy of the estimated demand curve depends on how well wage rates measure the opportunity cost of travel time; it will not be a good measure for people who view the travel itself as desirable.
Chapter 16: Cost–Benefit Analysis

Reprise of Monetization

Changes in social surplus serve as the basis for measuring the costs and benefits of policies. The concept of opportunity cost helps us value the inputs that policies divert from private use; the concept of willingness to pay helps us value policy outputs. The key to valuing outputs is to identify the primary markets in which they appear. When the outputs are not traded in organized markets, ingenuity is often needed to infer supply and demand schedules. Effects in undistorted secondary markets should be treated as relevant costs and benefits only if they involve price changes.

Discounting for Time and Risk

Policies often have effects far into the future. For example, deepening a harbor this year will allow large ships to use the harbor for perhaps a decade before dredging is again necessary. In determining the desirability of the harbor project, should we treat a dollar of benefit accruing next year as equivalent to a dollar of benefit accruing ten years from now? How should we take account of the possibility that dredging will be needed as soon as five years instead of the predicted ten? The concept of present value provides the basis for comparing costs and benefits that accrue at different times. The concept of expected value provides the commonly used approach to dealing with risky situations.

The Concept of Present Value

Most of us would be unwilling to lend someone $100 today in return for a promise of payment of $100 in a year’s time. We generally value $100 today more than the promise of $100 next year, even if we are certain that the promise will be carried out and that there will be no inflation. Perhaps $90 is the most that we would be willing to lend today in return for a promise of payment of $100 in a year’s time. If so, we say that the present value of receiving $100 next year is $90. We can think of the $90 as the future payment discounted back to the present. Discounting is the standard technique for making costs and benefits accruing at different times commensurate.

Imagine that you are the economics minister of a small country. The curve connecting points \( X_1 \) and \( X_2 \) in Figure 16.5 indicates the various combinations of current (period 1) and future (period 2) production that your country’s economy could achieve. By using available domestic resources to their fullest potential, your country could obtain any combination of production on the curve connecting \( X_1 \) and \( X_2 \). The locus of these combinations is what economists call the production possibility frontier for your economy. If all effort is put into current production, then \( X_1 \) can be produced and consumed in period 1, but production and consumption will fall to zero in period 2. Similarly, if all effort is put into preparing for production in period 2, then \( X_2 \) will be available in period 2, but there is no production in period 1. You undoubtedly find neither of these extremes attractive. Indeed, let us assume that you choose point \( i \) on the production possibility frontier as the combination of production and consumption in the two periods that you believe makes your country best off. The indifference curve labeled \( I_1 \) gives all the other hypothetical combinations of production and consumption in the two periods that you find equally satisfying to combination \((Q_1, Q_2)\). You consider any point to the northeast of this indifference curve better than any
Part IV: Doing Policy Analysis

point on it. Unfortunately, combination \((Q_1, Q_2)\) is the best you can do with domestic resources alone.

Now suppose foreign banks are willing to lend or borrow at an interest rate \(r\). That is, the banks are willing to either lend or borrow one dollar in period 1 in return for the promise of repayment of \((1+r)\) dollars in period 2. You realize that, with access to this capital market, you can expand your country’s consumption possibilities. For instance, if you choose point \(j\) on the production possibility frontier in Figure 16.5, then you could achieve a consumption combination anywhere along the line with slope \(-(1+r)\) going through point \(j\) by either borrowing or lending. This line is shown as connecting points \(R_1\) and \(R_2\), where \(R_2 = PV(C_1, C_2) = C_1 + C_2/(1+r)\). Each additional unit of consumption in period 1 costs \((1+r)\) units of consumption in period 2; each additional unit of consumption in period 2 costs \(1/(1+r)\) units of consumption in period 1. Once these intertemporal trades are possible, your country’s consumption possibility frontier expands from the original production possibility frontier connecting \(X_1\) and \(X_2\) to the discount line connecting \(R_1\) and \(R_2\).

The most favorable consumption possibility frontier results from choosing the production combination such that the discount rate line is just tangent to the
production possibility frontier. This occurs at point $j$. The most desirable consumption opportunity is shown at point $k$, which is on your indifference curve $I_2$, the one with the highest utility that can be reached along the consumption possibility frontier (the line connecting $R_1$ and $R_2$). Thus, your best policy is to set domestic production at $Z_1$ and domestic consumption at $C_1$ in period 1, lending the difference ($Z_1 - C_1$) so that in period 2 your country can consume $C_2$, consisting of $Z_2$ in domestic production and ($Z_2 - C_2$), which is equal to $(1 + r)$ times ($Z_1 - C_1$) in repaid loans.

Note that the most desirable point on the production possibility frontier gives us the largest possible value for $R_1$. But $R_1$ can be interpreted as the present value of consumption, which is defined as the highest level of consumption that can be obtained in the present period (period 1). We write the present value of consumption as $PV(C_1, C_2)$ and the present value of production as $PV(Z_1, Z_2)$. Each of these present values equals $R_1$.

We can arrive at an algebraic expression for the present value of consumption as follows: First, note that the present value of $C_1$ is just $C_1$. Second, note that the maximum amount that could be borrowed against $C_2$ for current consumption is $C_2 / (1 + r)$. Third, adding these two quantities together gives

$$PV(C_1, C_2) = C_1 + C_2 / (1 + r)$$

Similarly, we can express the present value of production as

$$PV(Z_1, Z_2) = Z_1 + Z_2 / (1 + r)$$

The present value formula is easily extended to more than two periods. For example, consider the present value of consumption in three periods, $PV(C_1, C_2, C_3)$. If we are in the second period, the present value of future consumption is $A_2 = C_2 + C_3 / (1 + r)$, the maximum quantity of consumption in the second period that can be obtained from $C_2$ and $C_3$. Now step back to the first period. The present value of consumption equals $C_1 + A_2 / (1 + r)$, which by substitution gives

$$PV(C_1, C_2, C_3) = C_1 + C_2 / (1 + r) + C_3 / (1 + r)^2$$

In general, we can write the present value for $N$ periods as

$$PV(C_1, C_2, \ldots, C_N) = C_1 + C_2 / (1 + r) + \cdots + C_N / (1 + r)^{N-1}$$

The key to understanding the use of present value in CBA is to recognize that at point $k$ in Figure 16.5 you are indifferent between having one additional unit of consumption in period 1 and $(1 + r)$ additional units in period 2. (The slope of $I_2$ at $k$, which gives your marginal rate of substitution between consumption in the two periods, equals the slope of the discount line.) Thus, in evaluating policies that make small changes to your consumption in the two periods, you should treat one dollar of change in period 1 as equivalent to $1 / (1 + r)$ dollars of change in period 2. This is equivalent to taking the present value of the changes.

In an economy with a perfect capital market, all consumers see the same interest rate so that they all have the same marginal rates of substitution between current and future consumption. Because everyone is willing to trade current and future consumption at the same rate, it is natural to interpret the interest rate as the appropriate social rate for discounting changes in social surplus occurring in different periods so that they can be added to measure changes in social welfare. Later in this chapter we discuss the appropriateness of interpreting market interest rates as social discount rates.
When doing CBA, we apply the discounting procedure by converting benefits and costs to present values. So, for instance, if benefits $B_t$ and costs $C_t$ occur $t$ periods beyond the current period, then their contribution to the present value of net benefits equals $(B_t - C_t)/(1 + d)^{t-1}$, where $d$ is the social discount rate. The present value of net benefits equals the sum of the discounted net benefits that occur in all periods.

We illustrate discounting with a simple example. Consider a city that uses a rural landfill to dispose of solid refuse. By adding larger trucks to the refuse fleet, the city would be able to save $100,000 in disposal costs during the first year after purchase and a similar sum each successive year of use. The trucks can be purchased for $500,000 and sold for $200,000 after four years. (The city expects to open a resource recovery plant in four years that will obviate the need for landfill disposal.) The city can currently borrow money at an interest rate of 10 percent. Should the city buy the trucks? Yes, if the present value of net benefits is positive.

To calculate the present value of net benefits, we must decide whether to work with real or nominal dollars—either one will lead to the same answer as long as we use it consistently.

Real, or constant, dollars control for changing purchasing power by adjusting for price inflation. When we compare incomes in different years, we typically use the purchasing power in some base year as a standard. For example, nominal per capita personal income in the United States was $3,945 in 1970 and $9,503 in 1980. If you have ever listened to an older person reminisce, you know that a dollar purchased more in 1970 than it did in 1980. The consumer price index (CPI) is the most commonly used measure of changing purchasing power. It is based on the cost of purchasing a standard market basket of goods. The cost in any year is expressed as the ratio of the cost of purchasing the basket in the current year to the cost of purchasing it in some base year. For instance, using 1972 as the base year, real per capita personal income was $4,265 in 1970 and $5,303 in 1980. Thus, while nominal per capita personal income rose over the decade by 141 percent, real per capita income rose by only 24 percent.

In CBA, we are looking into the future. Obviously, we cannot know precisely how price inflation will change the purchasing power of future dollars. Nevertheless, if we take the nominal market interest rate facing the city as its discount rate, then we must estimate future costs and benefits in inflated dollars. The reason is that the market interest rate incorporates the expected rate of inflation—lenders do not want to be repaid in inflated dollars. Thus one appropriate approach to discounting is to apply the nominal discount rate to future costs and benefits that are expressed in nominal dollars.

The column of Table 16.1 on the extreme right shows nominal benefits assuming a 4 percent annual rate of inflation. Assuming that the projected annual savings of $100,000 in years 2, 3, and 4 represent a simple extrapolation of the savings in the first year, and that the $200,000 liquidation benefit is based on the current price of used trucks, we can interpret these amounts as being expressed in constant dollars. (We are implicitly assuming that wage rates, gasoline prices, and other prices that figure into the benefit calculations increase at the same rate as the general price level.) To convert from real to nominal dollars, we simply multiply the amount in each year by $(1 + i)^M$, where $i$ is the assumed rate of inflation and $M$ equals the number of years beyond the current year that the costs and benefits accrue. For instance, the $100,000 of real savings in the fourth year would be $112,490 in nominal fourth-year dollars if the inflation rate held steady over the period at 4 percent. Using the 10 percent market interest rate seen by the city as the nominal discount rate, the present value of net benefits from purchasing the new trucks is $28,250. Thus, as long as no alternative equipment
configuration offering a greater present value of net benefits can be identified, the city should purchase the trucks.

An alternative method of discounting leads to the same present value of net benefits: apply a real discount rate to future costs and benefits that are expressed in constant (real) dollars. Projecting costs and benefits in constant dollars (that is, ignoring inflation) is natural. The difficulty arises in determining the appropriate discount rate. Nominal interest rates are directly observable in the marketplace; real interest rates are not. So when we decide that the interest rate facing the decision maker is the appropriate discount rate, we must adjust the observable nominal rates to arrive at real rates. This requires us to estimate the expected inflation rate just as we must if we decide to work with nominal costs and benefits. The real interest rate approximately equals the nominal interest rate minus the expected rate of inflation. More precisely:

\[
d = \frac{r - i}{1 + i}
\]

where \(d\) is the real discount rate, \(r\) is the nominal discount rate, and \(i\) is the expected rate of inflation.\(^{27}\)

For our city, which sees a nominal discount rate of 10 percent and expects a 4 percent inflation rate, the real discount rate is 5.77 percent. Applying this real discount rate to the yearly real costs and benefits in the lefthand column of Table 16.1 yields a present value of net benefits equal to $28,250. Thus, discounting real costs and benefits with the real discount rate is equivalent to discounting nominal costs and benefits with the nominal discount rate. As long as we use either real dollars and real

\[\text{Table 16.1 Present Value of the Net Benefits of Investment in New Garbage Trucks}\]

<table>
<thead>
<tr>
<th>Year 1:</th>
<th>Yearly Net Benefits Based on Real Dollars</th>
<th>Yearly Net Benefits Assuming 4 Percent Annual Inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purchase</td>
<td>-500,000</td>
<td>-500,000</td>
</tr>
<tr>
<td>Savings</td>
<td>100,000</td>
<td>100,000</td>
</tr>
<tr>
<td>Year 2:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savings</td>
<td>100,000</td>
<td>104,000</td>
</tr>
<tr>
<td>Year 3:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savings</td>
<td>100,000</td>
<td>108,160</td>
</tr>
<tr>
<td>Year 4:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savings</td>
<td>100,000</td>
<td>112,490</td>
</tr>
<tr>
<td>Year 5:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liquidation</td>
<td>200,000</td>
<td>233,970</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Present Value of Net Benefits</th>
<th>Real Discount Rate ((d = .0577))</th>
<th>Nominal Discount Rate ((r = .10))</th>
</tr>
</thead>
<tbody>
<tr>
<td>$28,250</td>
<td>$28,250</td>
<td>$28,250</td>
</tr>
</tbody>
</table>

Note: Let \(r\) = nominal discount rate, \(d\) = real discount rate, and \(i\) = expected inflation rate. Then \((1 + r) = (1 + d)(1 + i)\) or \(d = (r - i)/(1 + i)\). If \(r = .10\) and \(i = .04\), then \(d = .0577\).

\(^{27}\)The real discount rate equals the nominal rate less the expected inflation rate when discounting is done continuously rather than by discrete period. (Think of the distinction between annual and continuous compounding of interest.) In discounting by discrete period, we can separate the nominal discount factor \((1 + r)\) into the product of the real discount factor \((1 + d)\) and the constant dollar correction \((1 + i)\). From the expression \((1 + r) = (1 + d)(1 + i)\), we can solve for \(d = (r - i)/(1 + i)\).
discount rates or nominal dollars and nominal discount rates, we will arrive at the same present value.

The desirability of a policy often depends critically on the choice of the discount rate. Policies that involve building facilities, establishing organizations, or investing in human capital usually accrue costs before benefits: *when costs precede benefits, the lower the discount rate used, the greater is the present value of net benefits.* In the investment problem presented in Table 16.1, for example, the present value is over three times larger if a zero discount rate is used, and it becomes negative if a discount rate of over about 8.4 percent is used. Because analysts always have some uncertainty about the precise value of the appropriate discount rate, advocates of policies often argue for the lower discount rates than those who oppose them.

In situations in which there is likely to be controversy over the appropriate social discount rate, it is important to show the sensitivity of the present value of net benefits to the assumed social discount rate. When considering whether or not to adopt a particular project, it may be informative to search for a breakeven rate, the lowest discount rate at which the policy under consideration offers positive net benefits.

**Determining the Social Discount Rate.** We noted that lower discount rates generally make public investments appear more efficient. It should not be surprising, therefore, that the appropriate method for choosing the discount rate has been and continues to be hotly debated among theoretical and applied economists. Unfortunately, policy analysts do not have the luxury of waiting until a clear consensus develops.

In a world with perfect capital markets, illustrated in Figure 16.5, every consumer is willing to trade between marginal current and marginal future consumption at the market rate of interest. At the same time, the rate at which the private economy transforms marginal current consumption into marginal future consumption (the marginal rate of return on investment) also equals the market rate of interest. Thus, the market rate of interest would appear to be the appropriate social discount rate.

The situation becomes much more complicated when we relax our assumption of a perfect capital market. First, because individual consumers have finite lives, they may not sufficiently take into account the consumption possibilities of future generations. Those yet unborn do not have a direct voice in current markets, yet we may believe that they should have standing in our CBA. The willingness of people to pass inheritances to their children and to pay to preserve unique resources gives indirect standing to future generations. Furthermore, future generations inherit a growing stock of knowledge that will compensate at least somewhat for current consumption of natural resources. Nevertheless, to the extent that the interests of future generations remain underrepresented in current markets, an argument can be made for using a discount rate lower than the market interest rate.

The argument is particularly relevant when evaluating projects, like the storage of nuclear wastes, with consequences far into the future. Even very low positive discount rates make negligible the present value of costs and benefits occurring in the far future. Rather than trying to adjust the discount rate so that these far-off costs and benefits carry weight in a CBA, we believe that it is more appropriate to treat net benefits, discounted in the standard way, as the measure of one goal in a multigoal policy analysis.

---

Second, taxes and other distortions lead to divergence between the rate of return on private investment and the rate at which consumers are willing to trade current and future consumption. Suppose that consumers are willing to trade marginal current and marginal future consumption at a rate of 6 percent, their *marginal rate of pure time preference*.29 If they face an income tax of 25 percent and firms face a profits tax of 50 percent, then they will invest only in projects that earn a return of at least 16 percent. (The firm returns 8 percent to the investor after the profits tax is paid; the investor retains 6 percent after paying the income tax.)

Which, if either, of these rates should be interpreted as the social discount rate? Economists have generally answered this question in terms of the opportunity costs of the public project.30 If the public project is to be financed entirely at the expense of current consumption, then the marginal rate of pure time preference is the appropriate discount rate. If the public project is to be financed at the expense of private investment, then the rate of return on private investment is appropriate. The *shadow price of capital approach* allows for financing at the expense of both consumption and private investment by first converting reductions in private investment to the streams of benefits that they would have produced, and, once all financing is expressed in terms of lost consumption, discounting at the marginal rate of pure time preference.31

What can we do in the absence of a consensus on the appropriate social discount rate? One approach is to report net benefits for a range of discount rates. Analysts should also explain why they believe the range they have used is reasonable. A related approach is to report the largest discount rate that yields positive net benefits. Clients can then be advised to adopt the project if they believe the correct discount rate is smaller than the reported rate.

Another approach is to use the same discount rate for all projects being considered by the decision-making unit. With this approach, at least all projects are evaluated using the same standard. The U.S. Office of Management and Budget adopted this approach in 1972, requiring all federal agencies to use a real discount rate of 10 percent.32 Its 1992 guidelines lowered the generally used real discount rate to 7 percent.33 Congress exempted water projects, however, because many politically favored projects needed a lower discount rate to show positive net benefits—even CBA itself can be subverted by government failure! Of course, there may be losses rather than gains in efficiency from using a common standard if it is far from the correct value. Most economists would agree that even the 1992 OMB standard is too high. In contrast, the real discount rate used by the Congressional Budget Office, 2 percent, is probably a bit on the low side. A review of the issues and evidence suggest that a reasonable social discount rate for most projects without intergenerational impacts is about 3.5 percent.34

---

29We take this example from Lind, *Discounting for Time and Risk in Energy Policy*, 28–29.
30Ibid. 32–33.
Finally, there are some situations where market interest rates represent the opportunity cost of public investment and, therefore, are the appropriate discount rates. For example, if we are doing a CBA from the perspective of a local government, then the opportunity cost of public investment is the rate at which the local government can borrow. A similar argument can be made for small countries that borrow in international capital markets.

**The Concept of Expected Value**

Future costs and benefits can never be known with absolute certainty. Often, however, we know that certain future conditions, or contingencies, will influence costs and benefits. If we know which contingencies will arise, then we can make a much more accurate prediction. For example, if at least one major flood occurs in a river valley sometime over the next twenty years, then we might predict the present value of net benefits of building a dam to be, say, $25 million. If, on the other hand, a major flood does not occur, then we might predict the present value of net benefits of building the dam to be $−5 million. Of course, we do not know with certainty which of these contingencies will actually occur.

When dealing with contingencies in CBA, the standard approach is to assign probabilities to the various contingencies so that expected net benefits can be calculated. In common terminology, we convert the decision problem from one of uncertainty to one of risk by specifying contingencies and their probabilities of occurrence.

Once we have converted the problem to one of risk, we can apply the standard techniques of decision analysis. Specifically, we follow a four-step procedure. First, identify a set of mutually exclusive contingencies that cover all possibilities. For example, “one or more major floods in the next twenty years” and “no major floods in the next twenty years” are mutually exclusive and exhaust the possibilities. Second, for the policy being evaluated (the dam), estimate the present value of net benefits under each contingency ($25 million if one or more major floods, $−5 million if no major floods). Third, assign a probability of occurrence to each contingency such that the probabilities sum to 1. For instance, if over the last 100 years, there were two major floods in the river valley, then we might estimate the annual probability of a major flood occurring as .02, so that the probability of one or more floods in the next twenty years is .33; the probability of a major flood not occurring would be 1 minus the probability of a major flood occurring, or .67. Fourth, multiply the present value of net benefits for each contingency by the probability that the contingency will occur and sum to arrive at the expected value of present net benefits. To find the expected value of the present net benefits for the dam, for example, we evaluate the expression

\[
(.33)(25 \text{ million}) + (.67)(-5 \text{ million})
\]

which equals $4.9 million.

---


36The probability of no flooding in any year equals \(1 - .02 = 0.98\). Assuming that the probability of flooding in any year is independent of the probability in any other year, then the probability of at least one flood in the next twenty years is \(1 - .98^{20} = .33\).
In evaluating policies that have effects extending over a number of years, it is often appropriate to calculate the expected net benefits for each year and then discount to arrive at the present value of expected net benefits. For example, if we are evaluating a dam, we would almost certainly want our measure of net benefits to reflect when in the dam’s twenty-year useful life the major flood would occur. (Only if we believe that the value of avoided loss from a major flood grows at a yearly rate exactly equal to the discount rate would the timing of the flood be irrelevant.) To do this, we use our estimate of the annual probability of there being a major flood (.02—two flood years in the past century) so that we can calculate the expected value of net benefits for the \(i\)th year as:

\[
EN_{Bi} = (.02)NBF_i + (.98)NBNFi,
\]

where \(NBF_i\) equals the net benefits accruing in the \(i\)th year if a major flood occurs and \(NBNFi\) equals the net benefits accruing in the \(i\)th year if a major flood does not occur. The present value of expected net benefits for the dam is the sum of these yearly expected values discounted back to the present.

When we calculate expected net benefits on a yearly basis, we can allow for changing probabilities. For instance, if we believe that deforestation accompanying population growth in the river valley will gradually increase the chances of a major flood above the historical frequency, we can estimate yearly probabilities on the basis of either a hydrological model or some expert opinion and use them in our yearly calculations. Obviously, the more speculative the probabilities that we use, the more important it is that we test the sensitivity of the present value of expected net benefits to changes in the probabilities.

For most of us, the use of expected net benefits makes intuitive sense. Is its use conceptually valid as well? Returning to the guiding principle of willingness to pay, the conceptually correct measure of benefits is the sum of the amounts that all those affected by the project would be willing to pay for the project before they knew which contingency occurs. So, for example, the total benefit of a dam would be the sum of the amount that each farmer would be willing to pay before he or she knew whether there would be a drought or normal rain in the subsequent years. These amounts are called option prices.\(^{37}\) Although analysts sometimes attempt to elicit individuals’ option prices through contingent valuation surveys, they more commonly rely on expected values of outcomes over the possible contingencies. So, for example, valuing the benefits of the dam would be done by estimating the farmers’ producer surplus gains from the dam in a drought year, estimating their gains in non-drought years, and then finding an expected surplus applying the probabilities of drought and non-drought years to the surplus gains.

Option price can be either larger or smaller than expected surplus; the technical term for the difference is option value. In cases of collective risk, such as rainfall that affects the productivity of all farms in a river valley, option price is the conceptually correct measure of benefits; using expected surplus could result in either the adoption projects that do not increase efficiency or the rejection of projects that would have increased efficiency. Some theoretical progress has been made in determining the direction of option value, but only under fairly specific assumptions.\(^{38}\) In cases of individual risk, such as death from automobile accidents, the conceptually correct measure of benefits is the larger of option price and expected surplus; using expected surplus is a


A conservative approach to measuring benefits that could lead to the rejection of projects that would have increased efficiency.

Although expected surplus is not a perfect measure of benefits in cases of uncertainty, in most cases it will be close to the conceptually correct measure. As the estimation of costs and benefits from market data inevitably involves errors, the likely small errors associated with the use of expected surplus should not be immobilizing. In other words, the use of expected surplus is generally reasonable in practical CBA.

**Reprise of Discounting**

Most people value a dollar today more than the promise of a dollar tomorrow; they generally prefer a certain dollar more than the chance of receiving a dollar. Therefore, CBA requires us to discount for time and risk. Costs and benefits accruing in different time periods should be discounted to their present values. We can use either the real discount rate with constant dollars or the nominal discount rate with nominal dollars to arrive at present values. When we are able to express our uncertainty about costs and benefits in terms of contingencies and their probabilities, we can calculate expected net benefits. Adopting only policies with positive expected net benefits will generally lead to increased aggregate efficiency.

**Choosing among Policies**

Our discussion so far has largely considered the evaluation of single policies in isolation. If, after appropriately discounting for time and risk, a policy offers positive net benefits, then it satisfies the Kaldor-Hicks criterion and should be adopted (assuming, of course, that efficiency is the only relevant goal). A more general rule applies when we confront multiple policies that may enhance or interfere with each other: choose the combination of policies that maximizes net benefits. Physical, budgetary, and other constraints may limit the feasibility of generating such combinations.

**Physical and Budgetary Constraints**

Policies are sometimes mutually exclusive. For example, we cannot drain a swamp to create agricultural land and also preserve it as a wildlife refuge. When all of the available policies are mutually exclusive, we maximize efficiency by choosing the one with the largest positive net benefits. For example, consider the list of projects in Table 16.2. If we can choose any combination of projects, we simply choose all those with positive net benefits—namely, projects A, B, C, and D. Assume, however, that all the projects are mutually exclusive except C and D, which can be built together. By taking the combination of C and D to be a separate project, we can consider all the projects on the list to be mutually exclusive. Looking down the column labeled "Net Benefits," we see that project B offers the largest net benefits and therefore should be the one that we select.

Analysts sometimes compare programs in terms of their benefit-cost ratios. Note that project B, which offers the largest net benefits, does not have the largest ratio of benefits to costs. Project A has a benefit-cost ratio of 10, while project B has a benefit-cost ratio of only 3. Nevertheless, we select project B because it offers larger net benefits than project A. This comparison shows that the benefit-cost ratio can sometimes
confuse the choice process. Another disadvantage of the benefit-cost ratio is that it depends on how we take account of costs and benefits. For example, consider project B in Table 16.2. Suppose that the costs of $10 million consist of $5 million of public expenditure and $5 million of social surplus loss. We could take the $5 million of social surplus loss to be negative benefits so that costs would then be $5 million and benefits $25 million. While net benefits still equal $20 million, the benefit-cost ratio increases from 3 to 5. Therefore, we recommend that you avoid using benefit-cost ratios altogether.

Also note that projects C and D are shown as synergistic. That is, the net benefits from adopting both together exceed the sum of the net benefits from adopting each independently. Such might be the case if project C were a dam that created a reservoir that could be used for recreation and D were a road that increased access to the reservoir. Of course, projects can also interfere with each other. For instance, the dam might reduce the benefits of a downstream recreation project. The important point is that care must be taken to determine interactions among projects so that the combination of projects providing the greatest net benefits in aggregate can be readily identified.

Returning to Table 16.2, interpret the listed costs as public expenditures and the listed benefits as the monetized value of all other effects. Now assume that while none of the projects are mutually exclusive in a physical sense, total public expenditures (costs) cannot exceed $10 million. If project B is selected, the budget constraint is met and net benefits of $20 million result. If projects A, C, and D are selected instead, the budget constraint is also met but net benefits of $23 million result. No other feasible combination offers larger net benefits. Thus, under the budget constraint, choosing projects A, C, and D maximizes net benefits.

**Distributional Constraints**

The Kaldor-Hicks criterion requires only that policies have the potential to produce Pareto improvements; it does not require that people actually be compensated for the costs that they bear. One rationale for accepting potential, rather than demanding
actual, Pareto improvements for specific policies is that we expect different people to bear costs under different policies so that over the broad range of public activity few, if any, people will actually bear net costs. Another rationale is that even if some people do end up as net losers from the collection of policies selected on the basis of efficiency, they can be compensated through a program that redistributes income or wealth.

These rationales are less convincing for policies that concentrate high costs on small groups. If we believe that the losers will not be indirectly compensated, then we may wish to consider redesigning policies so that they either spread costs more evenly or provide direct compensation to big losers. We can think of limits on the size of losses as a constraint that must be met in applying the Kaldor-Hicks criterion. Alternatively, we can imagine applying weights to costs and benefits accruing to different groups—this would lead us to the modified cost–benefit analysis we discussed in Chapter 15. Using distributional constraints and weights obviously introduces values beyond efficiency.\(^{39}\) We should be careful to make these values explicit. Indeed, the best approach would be to treat net benefits as the measure of efficiency within a multigoal analysis.

Of course, the inclusion of distributional values requires that costs and benefits be disaggregated for relevant groups. This entails doing separate CBAs from the perspective of different income classes, geographic regions, ethnic groups, or whatever other categories that have distributional relevance. When you begin your analyses, err on the side of disaggregating too much. It may not be possible to gather necessary information to do distributional analysis at the end of a study. In addition, it may be that a distributional value will become important to you or your client only after you have seen the estimated distribution of net benefits. If efficiency does turn out to be the only relevant value, then aggregation is no more difficult than addition.

Disaggregation of net benefits by interest groups may be valuable in anticipating political opposition.\(^{40}\) While the estimation of aggregate net benefits enables us to answer the normative question—Is the policy efficient?—the estimation of net benefits by interest groups helps us answer the positive question—Who will oppose and who will support the policy? For example, Lee S. Friedman estimated net benefits from several perspectives in his evaluation of the supported work program in New York City.\(^{41}\) In addition to the social perspective, he estimated net benefits from the perspectives of taxpayers (Will the program continue to be politically feasible?), the welfare department (Will the program continue to be administratively feasible?), and a typical program participant (Will people continue to be willing to enter the program?). In this way, disaggregated CBA can serve as a first cut at predicting political and organizational feasibility.

**Reprise of Policy Choice**

When efficiency is the only relevant goal, we should choose the feasible combination of policies that maximizes net benefits. We should be especially careful in applying the net benefits rule when policies have interdependent effects. We should also anticipate

---

39 For a full discussion, see Boardman et al., *Cost–Benefit Analysis*, Chapter 18.
the possible introduction of distributional values by developing disaggregated net benefit estimates where practical. Finally, disaggregating net benefits by interest groups may enable us to anticipate political and organizational feasibility.

We now return to the example of the analysis of the alcohol tax discussed at the beginning of the chapter. It illustrates the application of many of the concepts introduced so far.

**An Illustration: Taxing Alcohol to Save Lives**

Alcohol is a widely used, and abused, substance. While some evidence suggests that moderate use of alcohol actually improves health, the medical literature reports that excessive use contributes to brain damage, cirrhosis of the liver, birth defects, heart disease, cancer of the liver, and a number of other adverse health problems. These alcohol-related conditions play at least some role in over 100,000 deaths per year in the United States. Yet, the most dramatic consequence of alcohol abuse is the large number of highway fatalities caused by alcohol-impaired drivers. In 1980, for example, more than 23,000 of the approximately 53,000 highway fatalities in the United States were caused by drivers who had been using alcohol. Young drivers, who have higher accident rates than adults, are especially dangerous when they have been drinking. Drivers under the age of 22 years are 100 times more likely to die in fatal crashes when they are under the influence of six or more drinks than when they are sober.

Not all of these adverse consequences of alcohol consumption fall on the drinker. On average, each 100 alcohol-impaired drivers who die in automobile accidents take with them about 77 victims. Nonfatal accidents caused by alcohol-impaired drivers inflict injury and property damage on nondrinkers. Publicly subsidized health insurance programs pay for some fraction of the costs of the morbidity caused by alcohol abuse. These external effects suggest that the market price of alcohol does not fully reflect its marginal social cost.

Recognizing these adverse external effects of alcohol consumption, and seeing a potential for generating substantial public revenue, a number of economists advocate that federal excise taxes on alcoholic beverages be raised. They note that per-unit taxes on beer and liquor have fallen greatly in real terms since the 1950s. The return of excise taxes to their previous real levels would not only raise revenue but would also internalize within the price of alcohol some portion of its external costs.

---

42This illustration was originally prepared in 1989, using the evidence then available. For an excellent contemporary assessment of alcohol control policies, see Philip J. Cook, *Paying the Tab: The Costs and Benefits of Alcohol Control* (Princeton, NJ: Princeton University Press, 2007).


Evaluating the economic efficiency of any particular tax increase requires us to identify and monetize the external effects of alcohol consumption. One of the more important effects is highway fatalities. How many highway fatalities will be avoided if people consume less alcohol because increased taxes raise the retail price? Can these lives be reasonably monetized? We concentrate on these questions in the context of a CBA of an increase in the excise tax on alcohol. Our analysis relies heavily on Charles Phelps’s estimates of the impact of tax increases on the highway fatality rates of young drivers.49

Estimating the Effects of a Tax on Alcohol

We consider the following impacts of an increase in alcohol taxes: social surplus losses in the alcohol markets (the major cost of the tax), reductions in fatalities caused by young and older drivers (benefits), reductions in the number of nonfatal highway accidents (benefits), and reductions in health and productivity losses (benefits). We discuss yearly effects with specific reference to a 30 percent tax on the retail prices of beer, wine, and liquor.

Social Surplus Losses in Alcohol Markets. In general, an excise tax reduces both consumer and producer surplus. If we assume that supply is perfectly elastic, however, the entire burden of the tax is borne by consumers. Figure 16.6 illustrates this

case for the beer market. The 1988 retail price of beer averaged about $0.63 per 12-ounce drink across the United States. At this price, 54 billion drinks are consumed annually in the United States. We assume that the beer industry would supply more or less at the same price; that is, the supply schedule is flat (perfectly elastic). If we impose a $0.19 tax per drink (30 percent of the current retail price), the retail price seen by consumers would rise to $0.82 per drink.

How will consumers respond? To answer this question, we must make assumptions about the demand schedule. Following the approach used by Phelps, we assume that demand is isoelastic. In particular, we assume that the demand schedule has the following form:

\[ q = a p^{-b} \]

where \( q \) is the quantity demanded, \( p \) is the price, and \( a \) and \( b \) are parameters. The elasticity of this demand schedule equals \(-b\), a constant. Relying on a review of the empirical literature, Phelps assumed that the price elasticity of demand for beer equals \(-0.5\) (a 1 percent increase in the price leads to a 0.5 percent reduction in the quantity demanded). As indicated in Figure 16.6, increasing the price to $0.82 under these assumptions reduces consumption to 47.4 billion drinks per year. The shaded area in the figure equals $9.538 billion, which is the loss in consumer surplus (equal to the sum of tax revenue and deadweight loss) in the beer market. The area of the shaded rectangle equals $8.951 billion, the portion of consumer surplus loss that is gained by the government as tax revenue. Thus, the net loss in social surplus in the beer market equals $0.586 billion, the difference between the consumer surplus loss and the government’s revenue gain. It is the deadweight loss we discussed in Chapter 4.

We follow the same procedure for estimating the social surplus losses in the wine and liquor markets. For the wine market, we assume that the price elasticity of demand equals \(-0.75\). With an average retail price of $0.46 per 5-ounce drink and current consumption of about 16 billion drinks per year, we find that a 30 percent tax on the current retail price yields a consumer surplus loss of $1.996 billion, a revenue gain of $1.814 billion, and a deadweight loss of $0.182 billion. Assuming an elasticity of \(-1.0\) in the liquor market, a 30 percent tax on the average retail price of $0.63 per drink results in a consumer surplus loss of $0.842 billion, a revenue gain of $0.754 billion, and a deadweight loss of $0.088 billion. The empirical literature suggests a demand elasticity of between \(-0.3\) and \(-0.4\). Wishing to be conservative in his estimations of the costs of tax increases, Phelps selected \(-0.5\) as his base case. (The higher the absolute value of the elasticity, the greater the consumer surplus loss in the beer market.) For a review of the empirical literature, see Stanley I. Ornstein, “Control of Alcohol through Price Increases,” *Journal of Studies on Alcohol*, 41, 1980, 807–18. Given an elasticity \((-b)\), the current price \((p)\), and the current quantity \((q)\), we can solve for the appropriate value of the constant \((a)\) in the demand schedule, \(q = a p^{-b}\).

Mathematically, this area equals the integral of the demand schedule between the initial price \((P_0)\) and the new price \((P_1)\). For our isoelastic demand schedule, the area equals

\[ \frac{a(z)}{z(P_1 - P_0)} \]

where \(z\) equals \((1 + b)\). If we were to assume a price elasticity of demand for beer of \(-2.5\) instead, we would find a social surplus loss of $0.306 billion. Assuming a demand elasticity of \(-0.75\) we would find a social surplus loss of $0.842 billion. When doing cost-benefit analyses, it is standard practice to test the sensitivity of our conclusions to changes in our assumptions. In other words, repeat the analysis, keeping all assumptions the same except the one under consideration. Given the consensus in the empirical literature, a range as large as \(-0.25\) to \(-0.75\) almost certainly covers the true value. We measure quantities in terms of drinks with approximately the same alcohol content: 12 ounces of beer, 5 ounces of wine, and 1.5 ounces of liquor.
1.5-ounce drink reduces annual consumption from 32 billion to 24.5 billion drinks and results in a consumer surplus loss of $5.205 billion, a revenue gain of $4.578 billion, and a deadweight loss of $0.627 billion.

Adding effects across the beer, wine, and liquor markets, we find that the 30 percent tax on retail prices reduces the consumption of alcoholic drinks by 17.6 percent, inflicts consumer surplus losses of $16.739 billion, generates tax revenue equal to $15.343 billion, and results in deadweight losses of $1.396 billion. We can either simply report the deadweight losses as the net cost of the policy in the alcohol markets or report the total consumer surplus losses as a cost and the total tax revenue as a benefit. Although the choice of approach will not alter our estimate of net benefits, the latter approach preserves information that may be relevant later if we decide that, along with economic efficiency, reducing the federal deficit is a relevant goal.

**Reductions in Fatalities Caused by Young Drivers.** The connection between alcohol taxes and fatal accidents caused by young drivers follows an intuitively straightforward chain: taxes raise the retail price of alcohol; higher prices lead young drivers to drink less; because they drink less, young drivers are involved in fewer fatal accidents. Michael Grossman, Douglas Coate, and Gregory M. Arluck estimated the price elasticity of youth drinking. Phelps combined this information with data from autopsies on the blood-alcohol levels of drivers killed in automobile accidents to estimate the effect of consumption changes on accident risks.

Grossman, Coate, and Arluck analyzed data from the U.S. Health and Nutrition Survey to estimate the effects of price, the legal drinking age, and other variables on the frequency and intensity of drinking by youths aged 16 to 21. They found that the frequency and intensity of drinking were highly sensitive to price. Indeed, they estimated that a 7 percent increase in the price of beer and liquor would have about the same effect as raising the minimum legal drinking age by one year. (Wine consumption appeared not to depend on price.) Their estimates permit the calculation of the probability of drinking level \( i \), \( P(\text{drinking level } i) \), and the probability of drinking frequency \( j \), \( P(\text{drinking frequency } j) \), for any set of assumed alcohol prices. The levels of drinking intensity indicate the number of drinks taken on a typical drinking day: 6 and over, 3 to 5, 1 to 2. The level of drinking frequency indicates how often youths drink: every day, 2 to 3 times per week, 1 to 4 times per month, 4 to 11 times per year, and never (fewer than 3 drinks in previous year).

In order to calculate the number of driver fatalities, we need the probability of a young driver dying in an automobile accident conditional on drinking level, \( P(\text{death }|\text{drinking level } i) \). That is, what is the probability that a youthful driver will be killed, given that he or she has had \( i \) drinks? Several states do autopsies, which measure blood-alcohol levels, on all highway fatalities. Data from these states permit the estimation of \( P(\text{drinking level } i|\text{death}) \), the probability that a young driver killed in a highway accident was drinking at level \( i \). Phelps realized that he could find the probability of death conditional on drinking level by using the following property of conditional probabilities:

\[
P(\text{death }|\text{drinking level } i)P(\text{drinking level } i) = P(\text{drinking level } i|\text{death})P(\text{death})
\]

55 National Highway Traffic Safety Administration, Appendix B.
Chapter 16: Cost–Benefit Analysis

where \( P(\text{death}) \) is the unconditional probability that a young driver will die in a highway accident. Because \( P(\text{death}) \) can be determined from aggregate statistics, it is possible to solve for the probability of death, given drinking level \( i \):

\[
P(\text{death} | \text{drinking level } i) = \frac{P(\text{drinking level } i | \text{death}) P(\text{death})}{P(\text{drinking level } i)}
\]

which is simply a statement of Bayes’ theorem.\(^{56}\) These probabilities change as the tax on alcohol changes because \( P(\text{drinking level } i) \) depends on the price of drinks.

Phelps next used the conditional probabilities of death and the probabilities of drinking frequencies estimated by Grossman, Coate, and Arluck to calculate the reduction in the number of young drivers who would die as the tax on beer is increased. Phelps concentrated on beer for two reasons: first, it is by far the drink of choice of youths; and, second, the price elasticity for beer estimated by Grossman, Coate, and Arluck was much more statistically precise than their estimates for wine and liquor. As long as the tax is actually imposed on wine and liquor as well as beer, Phelps’s approach leads to conservative estimates of the number of lives saved.

Phelps found that a 30 percent tax would result in 1,650 fewer drivers 16 to 21 years of age dying in highway fatalities per year. These driver fatalities are avoided because the tax decreases both the intensity and frequency of drinking. Because 77 non-driver fatalities are associated on average with each 100 driver fatalities, the 1,650 avoided driver deaths would result in an additional 1,270 lives saved per year. Thus, the 30 percent tax would reduce the number of highway fatalities caused by young drivers by about 2,920 per year.

**Reductions in Fatalities Caused by Older Drivers.** Unfortunately, no study comparable to that of Grossman, Coate, and Arluck is available for quantifying the behaviors of older drinkers in the United States. Consequently, we must use an ad hoc procedure that reflects two factors: first, on average, adults have a much less elastic demand for alcohol than do youths; and, second, for any given intensity of drinking, adults, on average, are less likely than youthful drivers to be involved in a fatal accident. As is often the case in CBA, we must do the best we can with available information.

\(^{56}\) A stylized example may be helpful to readers who have not encountered Bayes’ theorem before: Imagine that you are playing a game that involves guessing the proportion of red (\( R \)) and white (\( W \)) balls in an urn. Assume that you know that the urn contains either 8 \( W \) and 2 \( R \) balls (Type 1 urn) or 2 \( R \) and 8 \( W \) balls (Type 2 urn). Also assume that you believe that the prior probability of the urn being Type 1 is one-half (that is, before you sample the contents of the urn, you believe that \( P(\text{Type 1}) = P(\text{Type 2}) = .5 \)). Now imagine that you draw one ball from the urn and observe that it is white. What probability should you assign to the urn being type 1, given that you have observed a white ball? In other words, what is \( P(\text{Type 1} | W) \)? Bayes’ theorem tells you that

\[
P(\text{Type 1} | W) = \frac{P(W | \text{Type 1}) P(\text{Type 1})}{P(W)}
\]

But you know that

\[
P(W) = P(W | \text{Type 1}) P(\text{Type 1}) + P(W | \text{Type 2}) P(\text{Type 2}) = (8/10)(.5) + (2/10)(.5) = .5
\]

(In other words, because the urn types are equally likely before you sample, you can assume that you are drawing from their combined contents of 10 \( W \) and 10 \( R \) so that the probability of drawing a white ball is .5.) Thus, applying Bayes’ theorem gives

\[
P(\text{Type 1} | W) = \frac{(8/10)(.5)}{(8/10)(.5)} = .8
\]

Thus, if you draw a white ball, then you should believe that the probability of the urn being Type 1 is .8.
We start by noting that the 30 percent tax reduces the number of alcohol-involved fatalities of young drivers by about 40 percent (a reduction of about 1,650 on a base of about 4,120). Ignoring that adults have less elastic demand for alcohol and that alcohol-involved adult drivers are less dangerous, we might simply assume that the 30 percent tax would reduce the number of alcohol-involved fatalities of older drivers by the same fraction. With about 8,000 alcohol-involved fatalities each year among drivers over 21, the simple estimate would be 3,200. Obviously, this approach would lead to an overstatement of the number of fatalities involving drivers over 21 that would be avoided by the 30 percent tax.

In order to take account of the less elastic demand of adults, we might scale the simple estimate by the ratio of the adult elasticity of demand to the youth elasticity of demand. Using the Grossman, Coate, and Arluck results, Phelps calculated the elasticity of demand for beer for those 16 to 21 years old to be \(-2.3\). The elasticity of demand for older drinkers approximates the aggregate market elasticity because older drinkers constitute the large majority of the market. So the market elasticities we used in estimating effects in the alcohol markets (\(-0.5\) for beer, \(-0.75\) for wine, and \(-1.0\) for liquor) are reasonable approximations for the elasticity of drinkers older than 21. The ratio of the weighted average of these elasticities to the youth elasticity equals 0.3. Applying this ratio to the simple estimate of 3,240 adult drivers saved yields an adjusted estimate of 972 per year.

A lack of relevant data makes it more difficult to adjust further for the lower propensity of older drivers to be involved in fatal accidents at any given drinking level. We do know that while those 21 and younger constitute about 13 percent of licensed drivers, they account for about 26 percent of all alcohol-involved accidents.\(^{57}\) For lack of more relevant data, we assume that at any given drinking level, drivers over 21 are only 50 percent as likely to have fatal accidents as younger drivers. Adjusting for this factor reduces the estimate of adult drivers saved to 486. Applying the victim-to-driver ratio of 0.77 yields another 375 lives saved. Thus, we estimate the total number of lives saved due to changes in the behavior of drivers over 21 to be 861 per year.

**Reductions in Injuries and Property Damage.** Phelps estimated the injury and property damage costs of highway accidents involving alcohol-involved drivers to be about $3.75 billion per year, one-third of which is attributable to drivers 21 and younger.\(^{58}\) To estimate the injury and property damage savings associated with changes in youth drinking behavior, we assume that nonfatal accidents fall in the same proportion as fatal accidents. By multiplying the percentage reduction in alcohol-related deaths of young drivers resulting from a 30 percent tax (a 40 percent reduction) by the total costs of alcohol-involved nonfatal accidents ($1.25 billion) we obtain an annual savings of $0.5 billion. Following the same procedure for older drivers (a 0.06 reduction times $2.50 billion) yields annual savings of $0.15 billion. Thus we estimate the total annual injury and property damage savings from imposition of the 30 percent tax to be $0.65 billion.

---

\(^{57}\)National Highway Traffic Safety Administration, Appendix B, Table 6.

Health and Productivity Gains. We expect reductions in alcohol consumption to contribute to better health. We also expect reductions in alcohol consumption to increase productivity: better health and greater sobriety contribute to reductions in absenteeism and workplace accidents. For lack of more appropriate information, we assume that health and productivity losses are proportional to alcohol consumption.\(^59\) Therefore, because we expect the 30 percent tax to reduce alcohol consumption by 16.6 percent, we predict that health and productivity losses would also fall by 16.6 percent.

We must rely on previous studies for estimates of the yearly health and productivity costs associated with alcohol consumption. The best available estimates appear to be for the year 1975.\(^60\) Converting these estimates to 1986 dollars gives annual health costs equal to $25.92 billion and annual productivity costs equal to $39.96 billion. Therefore, under our assumption of proportionality between consumption and costs, we estimate $4.29 billion in annual health savings and $6.61 billion in annual productivity savings from the 30 percent tax.

Monetizing and Interpreting Effects

We have measured all effects in dollars except lives saved. We might think, therefore, that our only remaining task is to assign a dollar value to lives saved. Unfortunately, we must also question whether all the effects we have quantified belong in our calculation of net benefits. Do any of the effects involve double counting?

Which Lives Count? We estimated lives saved in four categories: young drivers, victims of young drivers, older drivers, and victims of older drivers. We can think of the victims of alcohol-involved drivers as suffering an externality of the drivers’ alcohol consumption. The costs borne by victims are not reflected in alcohol markets. In contrast, the costs borne by the drivers themselves may be reflected in their demand for alcohol. We expect that someone fully informed about the risks of driving under the influence of alcohol will consider these risks in deciding when and how much to drink. Other things equal, the higher the implicit values that drivers place on their own lives or the higher the probabilities of having fatal accidents after drinking, the less alcohol drivers will demand at any given price. To the extent that drivers are uninformed about the risks of driving under the influence of alcohol, their demand for alcohol will not fully reflect their risk of being an alcohol-involved driver fatality.

We can think of the alcohol markets as the primary markets and the “markets for victim and driver fatalities” as secondary markets. In the “fatality markets” people demand personal highway safety as a function of the “price” of safety, which can be thought of as the monetary equivalent of the level of effort expended on avoiding accidents. Now the market for victim fatalities is clearly distorted by the external effects of alcohol consumption. Therefore, we should count effects in the “victim market” in our CBA. The “driver market” is not distorted as long as drivers fully realize the increased risks they face from their alcohol consumption. If we believe that drivers are fully informed and not alcohol-addicted, then we should not count reductions in their

\(^59\) Proportional reduction would be unlikely if most of the reduction in alcohol consumption was by light drinkers rather than heavy drinkers. Evidence suggests, however, that even heavy drinkers do alter their behavior in the face of higher prices. See Hacker, “Taxing Booze for Health and Wealth.”

fatality rate as benefits—they are already counted in the alcohol markets. If we believe, however, that they do not fully take account of the increased risks, or that they are physically addicted to alcohol, then the “driver market” is distorted by an information asymmetry and we should count all or part of the avoided fatalities as benefits.

We estimate benefits under three different sets of assumptions about how well young and older drivers are informed about the risks of drinking and driving. First, we assume that all drinkers are uninformed (the belief of most parents!). Under this assumption, we count all avoided driver and victim fatalities as benefits. We can treat our estimate of benefits under this assumption as an upper bound on the true benefits. Second, we assume that all drinkers are fully informed. In this case, we count only avoided victim fatalities as benefits. This case provides a lower bound on benefits. Third, we make a “best guess” about the extent to which young and older drivers are informed. We assume that young drivers are only about 10 percent informed so that we count 90 percent of avoided young-driver fatalities as benefits. We also assume that older drivers are about 90 percent informed so that we count 10 percent of avoided older-driver fatalities as benefits. As in the other cases, we count all avoided victim fatalities as benefits.

Monetizing the Value of Life. What dollar value should we assign to avoided fatalities? This question may strike you as crass, especially if you think about putting a dollar value on the life of a specific person. Many of us would be willing to spend everything we have to save the life of someone we love. But we all implicitly put finite values on lives when we make decisions that affect risks to ourselves and those we love. Do you always buckle your seat belt? Do you make all your passengers buckle theirs? Do you have a smoke detector on every floor of your house? Do you have a working fire extinguisher? Do you always wear a helmet when riding your bicycle? Do you always drive within the legal speed limit? If you answer No to any of these questions, you are implicitly saying that you do not put an infinite value on life: you have decided to accept greater risks of fatality in order to avoid small certain costs.

The key distinction is between actual and statistical lives. Most of us are willing to spend great sums to save the lives of specific persons. For example, we spare no cost in trying to rescue trapped miners. Yet, we are less willing to take actions that reduce the probability of accidents; as a society we do not take all possible precautions to prevent miners from becoming trapped. Indeed, miners themselves sometimes knowingly accept higher risks by ignoring inconvenient safety rules. In other words, as long as we are dealing with probabilities rather than certainties, people seem willing to consider trade-offs between dollars and lives. By observing these trade-offs, we can impute a dollar value to a statistical life, the problem we face in our CBA of the alcohol tax.


62Phelps reports that the college students he surveyed underestimated the increased risks of driving after heavy drinking by a factor of more than 10. Charles E. Phelps, “Risk and Perceived Risk of Drunk Driving among Young Drivers,” *Journal of Policy Analysis and Management* 6(4) 1987, 708–14.

63The risk-premium approach is conceptually well grounded in the economic concept of willingness to pay. The other major approach, used by courts in deciding compensation in wrongful death cases, is to value life at the present value of forgone future earnings. While this approach provides abundant consulting opportunities for economists, it does not have as clear a conceptual basis in economic theory as the risk-premium approach and therefore should be avoided in CBA.
Many studies have attempted to measure how much people implicitly value their lives by seeing how much additional wage compensation they demand for working at riskier jobs. Imagine two jobs with identical characteristics except that one involves a 1/1,000 greater risk of fatal injury per year. If we observe people willing to take the riskier job for an additional $2,000 per year in wages, then we can infer that they are placing an implicit value on their (statistical) lives of $2,000/(1/1,000), or $2 million. The validity of our inference depends on the jobs' differing only in terms of risk and the workers fully understanding the risk.

In practice, researchers use econometric techniques to control for a wide range of job and worker characteristics. In his review of major studies of wage-risk trade-offs, W. Kip Viscusi found estimates of the value of life ranging from about $600,000 to over $8 million (in 1986 dollars). We take $1 million, near the lower end of this range, as a conservative estimate of the value of life.

**Apportioning Other Effects.** We must determine how much of the injury, property damage, health, and productivity effects to count as benefits under our uninformed demand, informed demand, and best guess cases. We follow the same line of argument that we used in deciding which avoided fatalities to count.

In the **uninformed demand** case, we assume that people do not consider the accident, health, and productivity costs of drinking so that we count the entire savings in these categories as benefits. This approach is consistent with treating the uninformed demand cases as an upper bound on benefits from the tax.

In the **informed demand** case, we assume that drinkers fully anticipate and bear these costs so that we do not count reductions in them as benefits. With respect to accident costs, for instance, we assume that the drinkers pay for the property damage and injuries they inflict on others through either higher insurance premiums or loss of coverage. Similarly, we assume that workers who are less productive because of their drinking bear most of the productivity losses in the form of lower wages. To the extent that insurance and wage rates do not fully reflect accident propensities, health risks, and productivity losses associated with drinking, our accounting will understate benefits—an approach consistent with treating the informed demand case as a lower bound on benefits.

In the **best guess** case, we apportion costs under the assumption that young drinkers are about 10 percent informed and older drinkers are about 90 percent informed of the health, productivity, and accident costs associated with their drinking. Because the majority of health and productivity losses accrue to older drinkers, we take 10 percent of the total savings in these categories as benefits. With respect to accidents, we count as benefits all the costs avoided by nondrivers as well as 90 percent of costs avoided by young drivers and 10 percent avoided by older drivers. Although lack of better information forces us to adopt these ad hoc assumptions, they probably provide a reasonable intermediate estimate of benefits.

---

Estimating Net Benefits

Having apportioned and valued effects, we are ready to estimate net benefits under each of the three cases. As long as we expect the same pattern of costs and benefits to persist over time, we need only look at the net benefits in the single year. If we think that costs and benefits will change substantially over time, then we should estimate net benefits for a number of years into the future and discount them back to the present. For instance, stricter enforcement of driving-while-intoxicated laws might reduce the frequency with which people drink and drive so that over time the number of lives saved from the tax will fall. When we expect costs and benefits to change over time, we should adopt a horizon equal in length to the period we expect the policy to be in effect.

Net Benefits of the 30 Percent Tax. Table 16.3 displays the costs and benefits of a 30 percent tax on the retail prices of beer, wine, and liquor. Note that the tax appears to offer positive net benefits in each of the three cases. We should be careful in our interpretation, however.

Table 16.3 Costs and Benefits of a 30 Percent Tax on Alcohol (billions of dollars)

<table>
<thead>
<tr>
<th></th>
<th>Uninformed Demand</th>
<th>Best Guess</th>
<th>Informed Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lives Saved ($1 million/life)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young drivers</td>
<td>1.65</td>
<td>1.49</td>
<td>0</td>
</tr>
<tr>
<td>Victims of young drivers</td>
<td>1.27</td>
<td>1.27</td>
<td>1.27</td>
</tr>
<tr>
<td>Older drivers</td>
<td>.49</td>
<td>.05</td>
<td>0</td>
</tr>
<tr>
<td>Victims of older drivers</td>
<td>.37</td>
<td>.37</td>
<td>.37</td>
</tr>
<tr>
<td>Subtotal</td>
<td>3.78</td>
<td>3.18</td>
<td>1.64</td>
</tr>
<tr>
<td>Injury and Property Damage Avoided</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Young drivers</td>
<td>.50</td>
<td>.47</td>
<td>0</td>
</tr>
<tr>
<td>Older drivers</td>
<td>.15</td>
<td>.02</td>
<td>0</td>
</tr>
<tr>
<td>Subtotal</td>
<td>.65</td>
<td>.49</td>
<td>0</td>
</tr>
<tr>
<td>Health and Productivity Costs Avoided</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health</td>
<td>4.29</td>
<td>.43</td>
<td>0</td>
</tr>
<tr>
<td>Productivity</td>
<td>6.61</td>
<td>.66</td>
<td>0</td>
</tr>
<tr>
<td>Subtotal</td>
<td>10.90</td>
<td>1.09</td>
<td>0</td>
</tr>
<tr>
<td>Tax Revenue</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beer, wine, and liquor</td>
<td>15.34</td>
<td>15.34</td>
<td>15.34</td>
</tr>
<tr>
<td>Consumer Surplus Change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beer, wine, and liquor</td>
<td>-16.74</td>
<td>-16.74</td>
<td>-16.74</td>
</tr>
<tr>
<td>Net Benefits</td>
<td>13.95</td>
<td>3.36</td>
<td>0.24</td>
</tr>
</tbody>
</table>
First consider the net benefits in the informed demand case. The reported net benefits of $0.24 billion are quite small when compared to the size of the costs and benefits. If we have underestimated consumer surplus losses by as little as 2 percent, then the true net benefits under the informed demand case would be negative instead of positive. Because we have been conservative in counting benefits in this case, it is probably reasonable for us to conclude that the tax at least breaks even. Nevertheless, the general point is that we should not put false confidence in the accuracy of our specific estimates.

Next consider the uninformed demand case. Here we report large positive net benefits of $13.95 billion annually. In inspecting the benefit categories, we note that almost 80 percent of the net benefits come from benefits under the heading of avoided health and productivity costs. Yet the estimates underlying these benefits were pulled somewhat uncritically out of the literature. Lack of time and access to primary data forced us to take at face value the yearly health and productivity costs estimated by other analysts. Our uncertainty about the accuracy of these estimates should give us concern. We may be victim to what has been called the problem of “horse and rabbit stew.” When we mix our fairly accurate estimates of avoided fatalities (rabbit) with the larger but less certain estimates of health and productivity savings (strong-flavored horse), our net benefits (stew) will be dominated by the less certain estimates (our stew will taste primarily of horse). The general point is that the uncertainty of the larger costs and benefits will largely determine the uncertainty of our net benefit estimate.

Fortunately, these problems are not as serious for the best guess case. Net benefits equal $3.36 billion, a fairly substantial amount even when compared to the large consumer surplus losses. Health and productivity benefits make up less than one-third of net benefits, reducing the danger that we have cooked horse and rabbit stew. Of course, whether or not our estimate of net benefits under the best guess is close to the true value of net benefits depends on the reasonableness of the various assumptions we have already discussed.

**Net Benefits of Other Tax Rates.** We focused on the 30 percent tax rate for purposes of exposition. Does the 30 percent rate offer the largest net benefits? Table 16.4

<table>
<thead>
<tr>
<th>Tax Rate (tax revenue)</th>
<th>Uninformed Demand</th>
<th>Best Guess</th>
<th>Informed Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>.10 (5.73)</td>
<td>6.24</td>
<td>2.09</td>
<td>0.63</td>
</tr>
<tr>
<td>.20 (10.80)</td>
<td>10.72</td>
<td>3.12</td>
<td>0.65</td>
</tr>
<tr>
<td>.30 (15.34)</td>
<td>13.95</td>
<td>3.36</td>
<td>0.24</td>
</tr>
<tr>
<td>.40 (19.46)</td>
<td>17.99</td>
<td>4.68</td>
<td>0.20</td>
</tr>
<tr>
<td>.50 (23.22)</td>
<td>21.60</td>
<td>5.79</td>
<td>0.03</td>
</tr>
</tbody>
</table>
shows net benefits for lower and higher rates for each of the three cases. Under the informed demand case, net benefits peak at the 20 percent tax rate. Thus, if we see this case as the appropriate one (because either we think it is the most likely or we wish to be conservative), then we should recommend 20 percent as the most efficient rate.

Under either the uninformed demand or best guess case, net benefits continue to rise up to 50 percent, the highest rate analyzed. This indicates that an even higher rate than 50 percent may be optimal. Two considerations, one methodological and the other substantive, suggest caution in this interpretation. First, the further we move from current policy, the less confidence we should have in our predictions of effects. Our assumption of isoeelastic demand, for instance, may be quite reasonable for small, but not large, price changes. Second, as we move toward prohibitive taxes, we may encounter radically different behavioral responses. Think of the ways people responded to Prohibition in the United States during the 1920s: they smuggled alcohol into the United States from other countries; they formed criminal organizations, which corrupted officials and practiced violence, to supply alcohol to the illegal market; they made alcohol at home; they developed a taste for more concentrated, and therefore more easily smuggled, forms of alcohol.67 Along with these sorts of behaviors, substantially higher taxes might induce greater use of other recreational drugs and perhaps trigger other important, but unanticipated, effects. Because we did not consider these possible effects, we should be cautious about advocating very high tax rates on the basis of our CBA alone.

Reconsidering the Value of Life: Switching to Cost-effectiveness. If we are unwilling to assign a dollar value to lives saved, then we must abandon CBA because we have incommensurable goals: saving lives and increasing economic efficiency. As long as we have only two goals, we can apply cost-effectiveness analysis (CEA) instead. We ask the basic question: How many dollars of economic efficiency must be given up for each life saved? For example, consider the informed demand case for the 30 percent tax. At a net cost of $1.4 billion, 1,645 lives can be saved, yielding a cost per life saved of $0.85 million.

We should compare the 30 percent tax rate with other tax rates in terms of cost per life saved. As the tax rate increases beyond 30 percent in the informed demand case, the number of lives saved goes up but so too does the cost per life saved. It is only when we compare alternatives that save the same number of lives that we can unequivocally say that the one with the lowest cost per life saved is best. Nevertheless, the cost per life saved indicates how high a dollar value of life we have to assume in a CBA to yield positive net benefits. As a rough comparison, a 1986 Office of Management and Budget review of health and safety regulations reported costs per life saved that ranged from approximately $100,000 to $72 million.68 A more recent review of more than 500 regulations and other actual or proposed policy interventions found that costs per life saved ranged from negative numbers (that is, the interventions actually saved resources as well as lives) to those costing over $10 billion (in 1993 dollars) per years of life saved.69

Valuing lives saved often arises in analysis of alternative health policies. Health researchers are often concerned not with prolonging life but also with the quality of the life enjoyed by those who live longer. They have developed a number of ways of estimating the values people place on various health states. Combining assessments of health states and duration of life leads to quality-adjusted life-years (QALYs), which are often used in cost-effectiveness analyses of alternative medical and health interventions.70

Reprise of Alcohol Tax CBA

Our analysis of alcohol taxation illustrates the basic craft and art of CBA. Basic concepts for measuring costs and benefits constitute the craft; drawing together fragmentary evidence from disparate sources to predict and monetize effects constitutes the art. Our conclusion that a 30 percent tax offers positive net benefits appears fairly robust to changes in assumptions about the measurement of benefits. Thus we could be fairly confident in concluding that it would increase economic efficiency if adopted. Yet we remain highly uncertain about whether 30 percent is close to the optimal rate.

Conclusion

CBA provides a framework for evaluating the economic efficiency of policies. The calculation of net benefits answers the question: Does the policy generate sufficient benefits so that those who bear its costs could at least potentially be compensated so that some people could be made better off without making anyone worse off? In order to calculate net benefits, we must decide which effects are relevant and how they can be made commensurable. The general concepts of opportunity cost and willingness to pay guide us in the application of the craft of CBA. The art lies in making reasonable inferences from data that are usually fragmentary and incomplete. The art also lies in realizing when inadequate data or social values other than efficiency make the narrow cost–benefit approach inappropriate.

For Discussion

1. Consider a regulation that would ban the use of cell phones while driving automobiles. List the effects of the regulation and classify them as costs and benefits.
2. Cost–benefit analysis provides a method for identifying the most efficient alternative. Consider its relationship to distributional values from both a theoretical and a practical perspective.