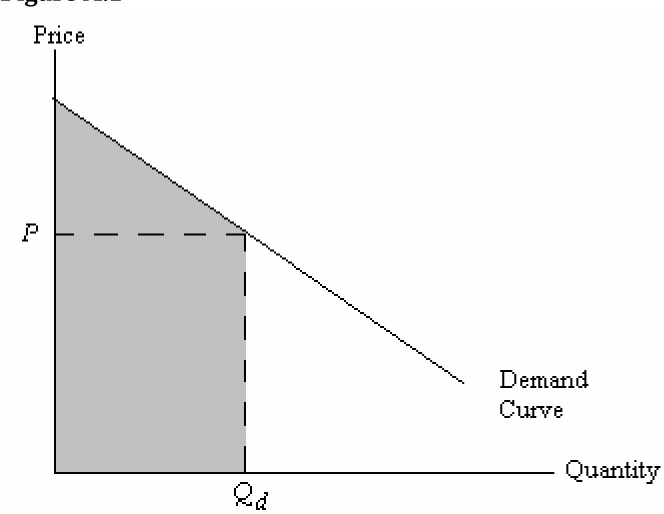
A. Economic Theory

This appendix provides a brief overview of the fundamental theory underlying the approaches to economic analysis discussed in Chapters 3 through 9. The first section summarizes the basic concepts of the forces governing a market economy in the absence of government intervention. Section A.2 describes why markets may behave inefficiently. If the preconditions for market efficiency are *not* met, government intervention can be justified.[250] The usefulness of benefit-cost analysis as a tool to help policy makers determine the appropriate policy response is discussed in Section A.3. Sections A.4 and A.5 explain how economists measure the economic impacts of a policy and set the optimal level of regulation. Section A.6 concludes and provides a list of additional references.

A.1 Market Economy

The economic concept of a market is used to describe any situation where exchange takes place between consumers and producers. Economists assume that consumers purchase the combination of goods that maximizes their well-being, or "utility", given market prices and subject to their household budget constraint, and that producers (firms) act to maximize their profits. Economic theory posits that consumers and producers are rational agents who make decisions taking into account *all* of the costs – the full opportunity costs[251] – of their choices, given their own resource constraints. The purpose of economic analysis is to understand how the agents interact and how their interactions add up to determine the allocation of society's resources: what is produced, how it is produced, for whom it is produced, and how these decisions are made. The simplest tool economists use to illustrate consumers' and producers' behavior is a market diagram with supply and demand curves.

Figure A.1



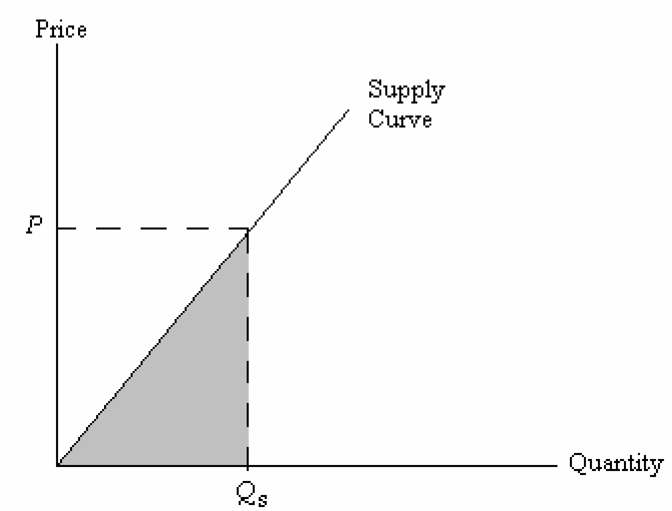
[250] EPA's mandates frequently rely on criteria other than economic efficiency as well, so policies are sometimes adopted that are not justified by the lack of efficiency.

[251] *Opportunity cost* is the next best alternative use of a resource. The full opportunity cost of producing (consuming) a good or service consists of the maximum value of other goods and services that could have been produced (consumed) had one not used the limited resources to produce (purchase) the good or service in question. For example, the full cost of driving to the store includes not only the price of gas but also the value of the time required to make the trip.

The demand curve for a single individual shows the quantity of a good or service that the individual will purchase at any given price (holding all else constant, i.e., assuming the budget constraint, information about the good, expected future prices, prices of other goods, etc. remain constant). The height of the curve indicates the maximum price, y, an individual with x units of a good or service would be willing to pay to acquire an additional unit of a good or service. This amount reflects the satisfaction (or utility) the individual receives from an additional unit, known as the *marginal benefit* of consuming the good. Economists generally assume that the marginal benefit of an additional unit is slightly less than that afforded by the previous unit so the amount an individual is willing to pay for one more unit of a good is less than the amount she paid for the last unit; hence, the individual demand curve slopes downward. A market demand curve shows the total quantity that consumers are willing to purchase at different price levels, i.e., their collective willingness-to-pay for the good or service. In other words, the market demand curve is the horizontal sum of all of the individual demand curves.

The concept of an individual's willingness to pay is one of the fundamental concepts used in economic analyses, and it is important to distinguish between total and marginal willingness to pay (WTP). Marginal WTP is the additional amount the individual would pay for one additional unit of the good. The total WTP is the aggregate amount the individual is willing to pay for the total quantity demanded (Q_d). Figure A.1 illustrates the difference between the marginal and total WTP. The height of the demand curve at a quantity Q_d gives the marginal WTP for the Q_d -th unit. The total WTP is equal to the marginal WTP for each unit up to Q_d – i.e., the shaded area under the demand curve from the origin up to Q_d.

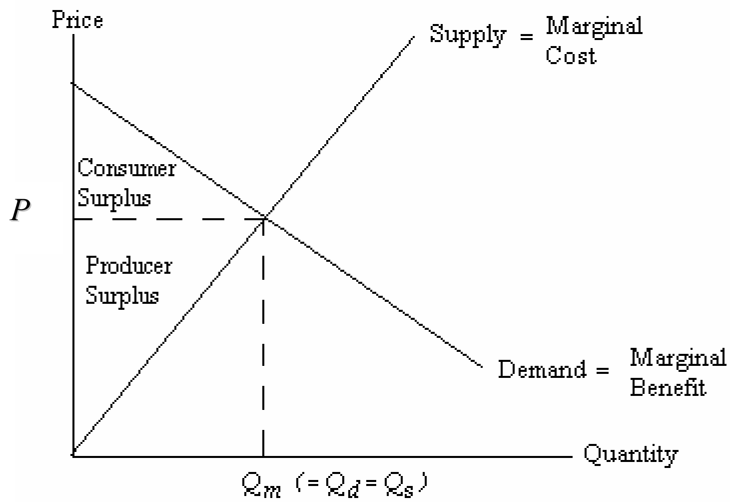
Figure A.2



An individual producer's supply curve shows the quantity of a good or service that an individual or firm is willing to sell (Q_s) at a given price. As a profit-maximizing agent, a producer will only be willing to sell another unit of the good if the market price is greater than or equal to the cost of producing that unit. The cost of producing the additional unit is known as the *marginal cost*. Therefore, the individual supply curve traces out the marginal cost of production and is also the marginal cost curve. Economists generally assume that the cost of producing one additional unit is greater than the cost of producing the previous unit because resources are scarce, and so the supply curve is assumed to slope upward. In Figure A.2, the marginal cost of producing the Q_s-th unit of the good is given by the height of the supply curve at Q_s. The *total cost* of producing Q_s units is equal to the shaded area under the supply curve from the origin to the quantity Q_s.[252] The market supply curve is simply the horizontal summation of the individual producers' marginal cost curves for the good or service in question.

[252] This is actually the long run total cost. In the short run there would be fixed costs as well.

Figure A.3



In a competitive market economy, the intersection of the market demand and market supply curves determines the equilibrium price and quantity of a good or service sold. The demand curve reflects the marginal benefit consumers receive from purchasing an extra unit of the good (i.e., it reflects their marginal willingness to pay for an extra unit). The supply curve reflects the marginal cost to the firm of producing an extra unit. Therefore, at the competitive equilibrium, the price is where the marginal benefit equals the marginal cost. This is illustrated in Figure A.3, where the supply curve intersects the demand curve at price P_m and quantity Q_m.

A counter-example illustrates why the equilibrium price and quantity occur at the intersection of the market demand and supply curves. In Figure A.3, consider some price greater than P_m where Q_s is greater than Q_d (i.e., there is *excess supply*). As producers discover that they cannot sell off their inventories, some may reduce prices slightly, hoping to attract more customers. At lower prices consumers will purchase more of the good (Q_d increases) although firms will be willing to sell less (Q_s decreases). This adjustment continues until Q_d equals Q_s. The reverse situation occurs if the price becomes lower than P_m. In that case, Q_d will exceed Q_s (i.e., there is *excess demand*) and consumers who cannot purchase as much as they would like are willing to pay higher prices. Therefore, firms will begin to increase prices, causing some reduction in the Q_d but also increasing Q_s. Prices will continue to rise until Q_s equals Q_d. At this point no purchaser or supplier will have an incentive to change the price or quantity; hence, the market is said to be in equilibrium.

Economists measure a consumer's net benefit from consuming a good or service as the excess amount that she is willing to spend on the good or service over and above the market price. The net benefit of all consumers is the sum of individual consumer's net benefits – i.e., what consumers are willing to spend on a good or service over and above that required by the market. This is called the *consumer surplus*. In Figure A.3, the market demands price P_m for the purchase of quantity Q_m. However, the demand curve shows that there are consumers willing to pay more than price P_m for all units prior to Q_m. Therefore, the consumer surplus is the area under the market demand (marginal benefit) curve but above the market price. Policies that affect market conditions in ways that decrease prices by decreasing costs of production (i.e., that shift the marginal cost curve to the right) will generally increase consumer surplus. This increase can be used to measure the benefits that consumers receive from the policy.[253]

[253] Section A.4.2 provides a more technical discussion of how consumer surplus serves as a measure of benefits.

On the supply side, a producer can be thought to receive a benefit if he can sell a good or service for more than the cost of producing an additional unit – i.e., its marginal cost. Figure A.3 shows that there are producers willing to sell up to Q_m units of the good for less then the market price, P_m. Hence, the net benefit to producers in this market, known as *producer surplus*, can be measured as the area above the market supply (marginal cost) curve but below the market price. Policies that increase prices by increasing market demand for a good (i.e., that shift the marginal benefit curve to the right) will generally increase producer surplus. This increase can be used to measure the benefits that producers receive from the policy.

Economic efficiency is defined as the maximization of social welfare. In other words, the efficient level of production is one that allows society to derive the largest possible net benefit from the market. This condition occurs where the (positive) difference between the total willingness to pay and total costs is the largest. In the absence of externalities and other market failures (explained below), this occurs precisely at the intersection of the market demand and supply curves where the marginal benefit equals the marginal cost. This is also the point where total surplus (consumer surplus + producer surplus) is maximized and there is no way to rearrange production or reallocate goods so that someone is made better off without making someone else worse off – a condition known as *Pareto optimality*. Notice that economic efficiency requires only that net benefits be maximized, *irrespective of to whom those net benefits accrue*. It does not guarantee an "equitable" or "fair" distribution of these surpluses among consumers and producers, or between sub-groups of consumers or producers.

Economists maintain that *if the economic conditions are such that there are no market imperfections* (as discussed in Section A.2), then this condition of Pareto optimal economic efficiency occurs automatically.[254] That is, no government intervention is necessary to maximize the sum of consumer surplus and producer surplus. This theory is summarized in the two Fundamental Theorems of Welfare Economics, which originate with Pareto (1906) and Barone (1908):

1. **First Fundamental Welfare Theorem.** Every competitive equilibrium is Pareto-optimal.
2. **Second Fundamental Welfare Theorem.** Every Pareto-optimal allocation can be achieved as a competitive equilibrium after a suitable redistribution of initial endowments.

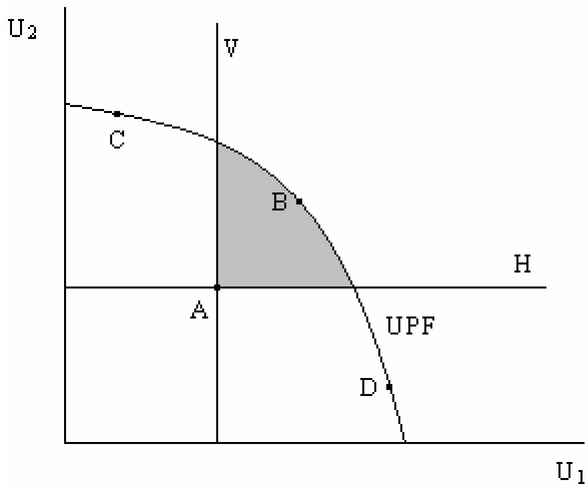
One graphical representation of these results is given in Figure A.4, which shows utility (welfare) levels in a two-person economy.[255] The curve shown is the utility possibility frontier (UPF) curve; the area within it represents the set of all possible welfare outcomes. Each point on the negatively sloped UPF curve is Pareto optimal since it is not possible to increase the utility of one person without decreasing the utility of the other. If the initial allocation is at point A, then the set of Pareto superior (welfare

[254] Technically, there are two types of efficiency. *Allocative efficiency* means that resources are used for the production of goods and services most wanted by society. *Productive efficiency* implies that the least costly production techniques are used to produce any mix of goods and services. Allocative efficiency requires that there be productive efficiency, but productive efficiency can occur without allocative efficiency. Goods can be produced at the least costly method without being most wanted by society. Perfectly competitive markets in the long run will achieve both of these conditions, producing the "right" goods (allocative efficiency) in the "right" way (productive efficiency). These two conditions imply Pareto optimal economic efficiency. (See Varian (1992) or any basic economics text for a more detailed discussion.)

[255] Another, perhaps more commonly used, graphical tool to explain the First and Second Welfare Theorems is an Edgeworth box. See Varian (1993) or other basic economic textbook for a detailed discussion.

enhancing) outcomes include all points in the shaded area, bordered by *H*, *V*, and the UPF curve.[256] If trading is permitted, the First Welfare Theorem applies and the market will move the economy to a superior, more efficient point such as *B*. Then the Second Welfare Theorem simply says that for any chosen point along the UPF curve, given a set of lump sum taxes and transfers, an initial allocation can be determined inside the UPF from which the market will achieve the desired outcome.[257]

Figure A.4



A.2 Reasons for Market or Institutional Failure

If the market supply and demand curves reflect society's true marginal social cost and willingness-to-pay, then a laissez-faire market (i.e., one governed by individual decisions and not government authority) will produce a socially efficient result. However, when markets do not fully represent social values, the private market will not achieve the efficient outcome (see Mankiw (2004), or any basic economics text); this is known as a *market failure*. Market failure is primarily the result of externalities, market power, and inadequate or asymmetric information. Externalities are the most likely cause of the failure of private and public sector institutions to account for environmental damages.

Externalities occur when markets do not account for the effect of one individual's decisions on another individual's well being.[258] In a free market, producers make their decisions about what and how much to produce taking into account the cost of the required inputs – labor, raw materials, machinery, energy – and consumers purchase goods and services taking into account their income and their own tastes and preferences. This means that decisions are based on the private costs and private benefits to market participants. If the consumption or production of these goods and services poses an external cost or

[256] Note that efficiency could be obtained by moving along the vertical line V, which keeps utility of person 1 constant while increasing utility of person 2, or by moving along the horizontal line H, which only shows improvements in utility for person 1. Moving to point B improves the utility for both individuals.

[257] Note that outcomes on the frontier such as C and D, although efficient, may not be desired on equity, or fairness, grounds.

[258] More formally, an externality occurs when the production or consumption decision of one party has an unintended negative (positive) impact on the profit or utility of a third party. Even if one party compensates the other party, an externality still exists.(Perman et al., 2003). See Baumol and Oates (1988) or any basic economics textbook for similar definitions and more detailed discussion.

benefit on those not participating in the market, however, then the market demand and supply curves no longer reflect the true marginal social benefit and marginal social cost. Hence, the market equilibrium will no longer be the socially (Pareto) efficient outcome.

Externalities can arise for many reasons. Transactions costs or poorly defined property rights can make it difficult for injured parties to bargain or use legal means to ensure that the costs of the damages caused by polluters are internalized into their decision making.[259] Activities that pose environmental risks may also be difficult to link to the resulting damages and often occur over long periods of time. Externalities involve goods that people care about but are not sold in markets.[260] Air pollution causes ill health, ecological damage, and visibility impacts over a long time period, and the damage is often far from the source(s) of the pollution. These additional social costs are not included in firms' profit maximization decisions and so are not considered when firms decide how much pollution to emit. Thus, the lack of a market for clean air causes problems and provides the impetus for government intervention in markets involving polluting industries.

Figure A.5 illustrates a negative externality associated with the production of a good. For example, a firm producing some product might also be generating pollution as a by-product. The pollution may impose significant costs – in the form of adverse health effects, for example – on households living downwind or downstream of the firm, but because those costs are not borne *by the firm,* the firm typically does not consider them in its production decisions. Society considers the pollution a cost of production, but the firm typically will not. In this figure:

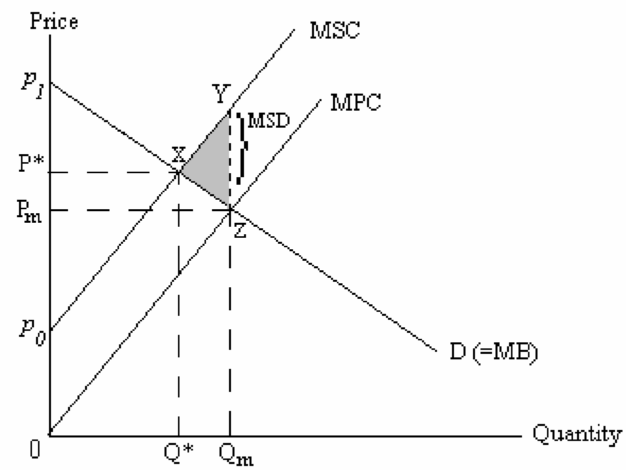
- D is the market demand (marginal benefit) curve for the product;
- MPC is the firm's marginal private real-resource cost of production, excluding the cost of the firm's pollution on households;
- MSD is the marginal social damage of pollution (or the marginal external cost) that the firm is not considering; and
- MSC is society's marginal social cost associated with production, including the cost of pollution (MSC = MPC + MSD).

---

[259] A property right can be defined as a bundle of characteristics that confer certain powers to the owner of the right: the exclusive right to the choice of use of a resource, the exclusive right to the services of a resource, and the right to exchange the resource at mutually agreeable terms. Externalities typically arise from the violation of one or more of the characteristics of well-defined property rights. This implies that the distortions resulting from an externality can be eliminated by appropriately establishing these rights. This insight is summarized by the famous "Coase theorem" which states that if property rights over an environmental asset are clearly defined, and bargaining among owners and prospective users of the asset is allowed, then externality problems can be corrected and the efficient outcome will result regardless of who was initially given the property right. (The seminal paper is Coase (1960).)

[260] Often these are goods that exhibit public good characteristics. Pure public goods are those which are non-rivalrous in consumption and non-excludable. (See Perman et al. (2003) for a detailed discussion of these, as well as congestible and open access resources — i.e., goods that are neither pure public nor pure private goods.) Because exclusive property rights cannot be defined for these types of goods, pure private markets cannot provide for them efficiently.

Figure A. 5



In an incomplete market, producers pay no attention to external costs, and production occurs where market demand and the marginal private real-resource cost (MPC) curves intersect – at a price P_m and a quantity Q_m. In this case, net social welfare (total willingness to pay minus total social costs) is equal to the area of the triangle p_0p_1X less the area of triangle *XYZ*.[261] If the full social cost of production, including the cost of pollution, is taken into consideration, then the marginal cost curve should be increased by the amount of the marginal social damage (MSD) of pollution. Production will now occur where the demand and marginal social cost (MSC) curves intersect – at a price *P** and a quantity *Q**. At this point net social welfare (now equal to the area of the triangle, p_0p_1X alone) is maximized, and therefore the market is at the socially efficient point of production. This example shows that when there is a negative externality such as pollution, and the social damage (external cost) of that pollution is not taken into consideration, the producer will oversupply the polluting good.[262] The shaded triangle (*XYZ*), referred to as the *deadweight loss*, represents the amount that society loses by producing too much of the good.

A.3 Benefit-Cost Analysis

If a negative externality such as pollution exists, an unregulated market will not account for its cost to society, and the result will be an inefficient outcome. In this case, there may be a need for government intervention to correct the market failure. A correction may take the form of dictating the allowable level of pollution or introducing a market mechanism to induce the optimal level of pollution.[263] Figure A.5 neatly summarized this in a single market diagram. To estimate the *total* costs and benefits to society of an activity or program, the costs and benefits in each affected market, as well as any non-market costs or benefits, are added up. This is done through Benefit-Cost Analysis (BCA).

BCA can be thought of as an accounting framework of the overall social welfare of a program, which illuminates the tradeoffs involved in making different social investments (Arrow et al., 1996). It is used

[261] Recall from Section A.1 that total willingness to pay is equal to the area under the demand curve from the origin to the point of production ($0p_1ZQ_m$). Total costs (to society) are equal to the area under the marginal social cost curve (MSC) from the origin to the point of production ($0p_0YQ_m$).

[262] Similarly, the private market will undersupply goods for which there are positive externalities, such as parks and open space.

[263] Chapter 4 discusses the various regulatory techniques and some non-regulatory means of achieving pollution control.

to evaluate the favorable effects of a policy action and the associated opportunity costs. The favorable effects of a regulation are the benefits, and the foregone opportunities or losses in utility are the costs. Subtracting the total costs from the total monetized benefits provides an estimate of the regulation's net benefits to society. An efficient regulation is one that yields the maximum net benefit, assuming that the benefits can be measured in monetary terms.

Benefit-cost analysis can also be see as a type of market test for environmental protection. In the private market, a commodity is supplied if the benefits that society gains from its provision, measured by what consumers are willing to pay, outweigh the private costs of producing the commodity. Economic efficiency is measured in a private market as the difference between what consumers are willing to pay for a good and what it costs to produce it. Since clean air and clean water are public goods, private suppliers cannot capture their value and sell it. The government determines their provision through environmental protection regulation. BCA quantifies the benefits and costs of producing this environmental protection in the same way as the private market, by quantifying the willingness to pay for the environmental commodity. As with private markets, the efficient outcome is the option that maximizes net benefits.

As mentioned above, the key to performing BCA lies in the ability to measure both benefits and costs in monetary terms so that they are comparable. The consumers and producers in regulated industries and the governmental agencies responsible for implementing and enforcing the regulation (and by extension, taxpayers in general) typically pay the costs. The total cost of the regulation is found by summing the costs to these individual sectors. (An example of this, excluding the costs to the government, is given in Section A.4.3.) Since environmental regulation usually addresses some externality, the benefits of the regulation often occur outside of markets. For example, the primary benefits of drinking water regulations are improvements in human health. Once the expected reduction in illness and premature mortality associated with the regulation has been calculated, economists use a number of techniques to estimate the value that society places on these health improvements.[264] These monetized benefits can then be summed to obtain the total benefits from the regulation.

Note that, in BCA, gains and losses are weighted equally regardless of to whom they accrue. Evaluation of the fairness, or the equity, of the net gains cannot be made without specifying a social welfare function. However, there is no generally agreed-upon social welfare function, and assigning relative weights to the utility of different individuals is an ethical matter that economists strive to avoid. Given this dilemma, economists have tried to develop criteria for comparing alternative allocations where there are winners and losers without involving explicit reference to a social welfare function. According to the Kaldor-Hicks compensation test, named after its originators Nicholas Kaldor and J.R. Hicks, a reallocation is a welfare-enhancing improvement to society if:

1. The winners could theoretically compensate the losers and still be better off, and
2. The losers could not, in turn, pay the winners to not have this reallocation and still be as well off as they would have been if it did occur (Perman et al. 2003).

While these conditions sound complex, they are met in practice by assessing the net benefits of a regulation through BCA. The policy that yields the highest positive net benefit is considered welfare enhancing according to the Kaldor-Hicks criterion. Note that the compensation test is stated in terms of *potential* compensation and does not solve the problem of evaluating the fairness of the distribution of well-being in society. Whether and how the beneficiaries of a regulation should compensate the losers involves a value judgment and is a separate decision for government to make.

[264] Chapter 7 discusses a variety of methods economists use to value environmental improvements.

Finally, BCA may not provide the only criterion used to decide if a regulation is in society's best interest. There are often other, overriding considerations for promulgating regulation. Statutory instructions, political concerns, institutional and technical feasibility, enforceability, and sustainability are all important considerations in environmental regulation. In some cases, a policy may be considered desirable even if the benefits to society do not outweigh its costs, particularly if there are ethical or equity concerns.[265] There are also practical limitations to BCA. Most importantly, it requires assigning monetized values to non-market benefits and costs. In practice, it may be very difficult or impossible to quantify gains and losses in monetary terms (e.g., the loss of a species, intangible effects).[266] In general, however, economists believe that BCA provides a systematic framework for comparing the social costs and benefits of proposed regulations, and that it contributes useful information to the decision-making process about how scarce resources can be put to the best social use.

A.4 Measuring Economic Impacts

A.4.1 Elasticities

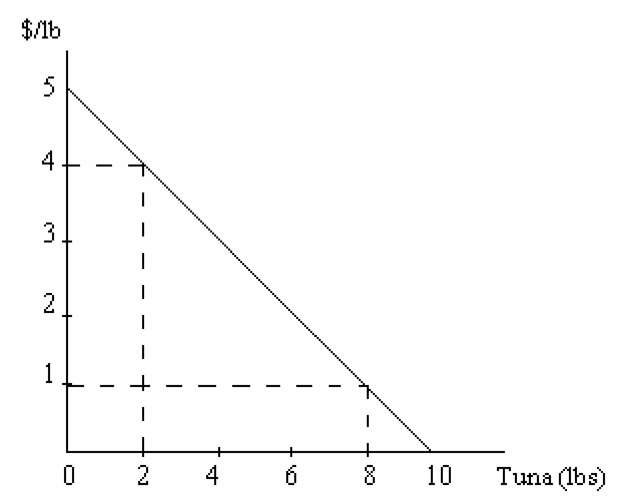
The net change in social welfare brought about by a new environmental regulation is the sum of the negative effects (i.e., loss of producer and consumer surplus) and the positive effects (or social benefits) of the improved environmental quality. This is shown graphically for a single market in Figure A.5 above. The use of demand and supply curves highlights the importance of assessing how individuals will respond to changes in market conditions. The net benefits of a policy will depend on how responsive producers and consumers' decisions are to a change in price. Economists measure this responsiveness by the supply and demand elasticities.

The term "elasticity" refers to the sensitivity of one variable to changes in another variable. The price elasticity of demand (or supply) for a good or service is equal to the percentage change in the quantity demanded (or supplied) that would result from a one percent increase in the price of that good or service. For example, a price elasticity of demand for tuna equal to -1 means that a 1% increase in the price of tuna results in a 1% decrease in the quantity demanded. Changes are measured assuming all other things, such as incomes and tastes, remain constant. Demand and supply elasticities are rarely constant and often change depending on the quantity of the good consumed or produced. For example, according to the demand curve for tuna shown in Figure A.6, at a price of $1 per pound, a 10% increase in price would reduce quantity demanded by 2.5% (from 8 lbs to 7.8 lbs). At a price of $4 per pound, a 10% increase in price would result in a 40% decrease in quantity demanded (from 2 to 1.2 lbs). This implies that the price elasticity of demand is -0.25 when tuna costs $1/lb but -4 when the price is $4/lb. Therefore, when calculating elasticities it is important to state the price or quantity of the good demanded (or supplied).

[265] Chapter 9 addresses equity assessment and describes the methods available for examining the distributional effects of a regulation.

[266] Kelman (1981) argues that it is even unethical to try to assign quantitative values to non-marketed benefits.

Figure A. 6



Elasticities are important in measuring economic impacts because they determine how much of a price increase will be passed on to the consumer. For example if a pollution control policy leads to an increase in the price of a good, multiplying the price increase by current quantity sold generally will not provide an accurate measure of impact of the policy. Some of the impact will take the form of higher prices for the consumer, but some of the impact will be a decrease in the quantity sold. The amount of the price increase that is passed on to consumers is determined by the elasticity of demand relative to supply (as well as existing price controls). "Elastic" demand (or supply) indicates that a small percentage increase in price results in a larger percentage decrease (increase) in quantity demanded (supplied).[267] All else equal, an industry facing a relatively elastic demand is less likely to pass on costs to the consumer because increasing prices will result in reduced revenues. Supply characteristics in the industries affected by a regulation can be as important as demand characteristics in determining the economic impacts of a rule. For highly elastic supply curves relative to the demand curves, it is likely that cost increases or decreases will be passed on to consumers.

The many variables that affect the elasticity of demand include:

- The cost and availability of close substitutes;
- The percentage of income a consumer spends on the good;
- How necessary the good is for the consumer;
- The amount of time available to the consumer to locate substitutes;
- The expected future price of the good; and
- The level of aggregation used in the study to estimate the elasticity.

The availability of close substitutes is one of the most important factors that determine demand elasticity. A product with close substitutes at similar prices tends to have an elastic demand, because consumers can readily switch to substitutes rather than paying a higher price. Therefore, a company is less likely to be able to pass through costs if there are many close substitutes for its product. Narrowly defined markets (e.g., salmon) will have more elastic demands than broadly defined markets (e.g., food) since there are more substitutes for narrow goods.

---

[267] Demand (or supply) is said to be "elastic" if the absolute value of the price elasticity of demand (supply) is greater than one and "inelastic" if the absolute value of the elasticity is less than one. If a percentage change in price leads to an equal percentage change in quantity demanded (supplied) (i.e., if the absolute value of elasticity equals one), demand (supply) is "unit elastic".

Whether the affected product represents a substantial or necessary portion of customers' costs or budgets is another factor that affects demand elasticities. Goods that account for a substantial portion of consumers' budgets or disposable income tend to be relatively price elastic. This is because consumers are more aware of small changes in the price of expensive goods compared to small changes in the price of inexpensive goods, and therefore may be more likely to seek alternatives. A similar issue concerns the type of final good involved. Reductions in demand may be more likely to occur when prices increase for "luxuries" or optional purchases than for basic requirements. If the good is a necessity item, the quantity demanded is unlikely to change drastically for a given change in price and demand will be relatively inelastic.

Elasticities tend to increase over time, as firms and customers have more time to respond to changes in prices. Although a company may face an inelastic demand curve in the short run, it could experience greater losses in sales from a price increase in the long run, as customers begin to find substitutes or as new substitutes are developed. However, temporary price changes may affect consumers' decisions differently than permanent ones. The response of quantity demanded during a 1-day sale, for example, will be much greater than the response of quantity demanded when prices are expected to decrease permanently. Finally, it also is important to keep in mind that elasticities differ at the firm versus the industry level. It is not appropriate to use an industry-level elasticity to estimate the ability of only one firm to pass on compliance costs when its competitors are not subject to the same cost.

Characteristics of supply in the industries affected by a regulation can be as important as demand characteristics in determining the economic impacts of a rule. For relatively elastic supply curves, it is likely that cost increases or decreases will be passed on to consumers. The elasticity of supply depends, in part, on how quickly costs per unit rise as firms increase their output. Among the many variables that influence this rise in cost are:

- The cost and availability of close input substitutes;
- The amount of time available to adjust production to changing conditions;
- The degree of market concentration among producers;
- The expected future price of the product;
- The price of related inputs and related outputs; and
- The speed of technological advances in production that can lower costs.

Similar to the determinants of demand elasticity, the factors influencing the price elasticity of supply all relate to a firm's degree of flexibility in adjusting production decisions in response to changing market conditions. The more easily a firm can adjust production levels, find input substitutes, or adopt new production technologies, the more elastic is supply. Supply elasticities tend to increase over time as firms have more opportunities to renegotiate contracts and change production technologies. When production takes time, the quantity supplied may also be more responsive to expected future price changes than to current price changes.

Demand and supply elasticities are available for the aggregate output of final goods in most industries. They are usually published in journal articles on research pertaining to a particular industry.[268] When

---

[268] Another useful source of elasticity estimates is the recently developed EPA Elasticity Databank. In the absence of an encyclopedic 'Book of Elasticities' the Elasticity Databank serves as a searchable database of elasticity parameters across a variety of types (i.e., demand and supply elasticities, substitution elasticities, income elasticities, and trade elasticities) and economic sectors/product markets. The database is populated with EPA generated estimates used in Environmental Impact Assessment (EIA) studies conducted by the Agency since

such information is unavailable, as is often the case for intermediate goods, elasticities may be quantitatively or qualitatively assessed.[269] Econometric tools are frequently used to estimate supply and demand equations (thereby the elasticities) and the factors that influence them.

A.4.2 Measuring the welfare effect of a change in environmental goods

As introduced in Section A.1, changes in consumer surplus are measured by the trapezoidal region below the ordinary, or Marshallian, demand curve as price changes. This region reflects the benefit a consumer receives by being able to consume more at a lower price. If the price of a good decreases, some of the consumer's satisfaction comes from being able to consume more of a commodity when its price falls, but some of it comes from the fact that the lower price means that the consumer has more income to spend. However, the change in (Marshallian) consumer surplus only serves as a monetary measure of the welfare gain or loss experienced by the consumer under the strict assumption that the marginal utility of income is constant.[270] This assumption is almost never true in reality. Luckily, there are alternative, less demanding monetary measures of consumer welfare that prove useful in treatments of benefit-cost analysis. Intuitively, these measures determine the size of payment that would be necessary to compensate the consumer for the price change. In other words, they estimate the consumer's WTP for a price change.

As mentioned above, a price decline results in two effects on consumption. The change in relative prices will increase consumption of the cheaper good (the substitution effect), and consumption will be affected by the change in overall purchasing power (the income effect). A Marshallian demand curve reflects both substitution and income effects; movements along it show how the quantity demanded changes as price changes (holding all other prices and income constant), so it reflects both the substitution and the income effects. The Hicksian (or "compensated") demand curve, on the other hand, shows the relationship between quantity demanded of a commodity and its price, holding all other prices and *utility* (rather than income) constant. This is the correct measure of a consumer's WTP for a price change. The Hicksian demand curve is constructed by adjusting income as the price changes so as to keep the consumer's utility the same at each point on the curve. In this way, the income effect of a price change is eliminated and movements along the Hicksian demand function can be used to determine the monetary change that would compensate the consumer for the price change, considering the substitution effect alone.

Hicks (1941) developed two correct monetary measures of utility change associated with a price change: compensating variation and equivalent variation. *Compensating variation* (CV) assesses how much money must be taken away from consumers after a price decrease occurred to return them to the original utility level. It is equal to the amount of money that would 'compensate' the consumer for the price decrease. *Equivalent variation* (EV) measures how much money would need to be given to the consumer to bring her to the higher utility level instead of introducing the price change. In other words, it is the monetary change that would be 'equivalent' to the proposed price change.

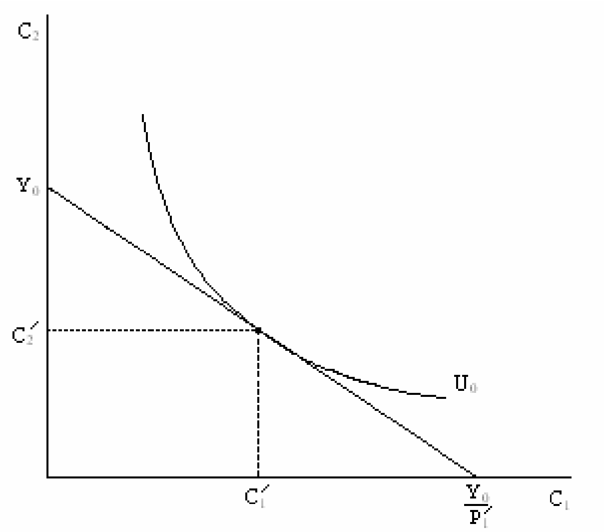
1990 as well as estimates found in the economics literature. It may be accessed from the Technology Transfer Network Economics & Cost Analysis Support website: http://www.epa.gov/ttnecas1/Elasticity.htm.

[269] Final goods are those that are available for direct use by consumers and are not utilized as inputs by firms in the process of production. Goods that contribute to the production of a final good are called intermediate goods. It is of course possible for a good to be final from one perspective and intermediate from another (Pearce, 1992).

[270] See Perman et al. (2003), Just et al. (2005), or any graduate level text for a more thorough exposition of this issue.

Before examining the implications of these measures for valuing environmental changes, it is useful to understand CV and EV in the case of a reduction in the price of some normal, private good, C_1.[271] This can be shown with indifference curves and a budget line, as seen in Figure A.7.

Figure A. 7



Assume that the consumer is considering the tradeoff between C_1 and all other goods, denoted by a composite good, C_2. The indifference curve, U_0, depicts the different combinations of the two goods that yield the same level of utility. Because of diminishing marginal utility, the curve is concave, where increasing amounts of C_1 must be offered for each unit of C_2 given up to keep the consumer indifferent. The budget line on the graph reflects what the consumer is able to purchase given her income, Y_0, and the prices of the two goods— P_1' and P_2', respectively.[272] A utility-maximizing consumer will choose quantities C_1' and C_2', the point where the indifference curve is tangent to the budget constraint.[273]

Figure A.8 shows the change in the optimal consumption bundle resulting from a reduction in the price of C_1. If the price of C_1 falls, the budget line shifts out on the C_1 axis because more C_1 can be purchased for a given amount of money. The consumer now chooses C_1'' and C_2'' at point *b* and moves to a new, higher utility curve, U_1. CV then measures how much money must be taken away at the new prices to return the consumer to the old utility level. That is, starting at point b and keeping the slope of the budget line fixed at the new level, by how much must it be shifted downward to make it tangent to the initial indifference curve, U_0? It is, therefore, the maximum amount the consumer would be willing to pay to have the price fall occur—i.e., the precise monetary measure of the welfare change[274] In Figure A.8, CV is simply given by the amount $Y_0 - Y_1$. EV, on the other hand, measures how much income must be given to the individual at the old price set to maintain the same level of well-being as if the price change did occur. That is, keeping the slope of the budget line fixed at the old level, by how much must it be shifted upwards to make it tangent to U_1? EV is, then, the minimum amount of money the consumer would accept in lieu of the price fall. This too is a proper monetary measure of the utility change resulting from the price decrease. In Figure A.8 then EV is the amount $Y_2 - Y_0$, leaving the individual at point *f*.

[271] The notation and discussion in this section follow Chapter 12 of Perman et al. (2003).

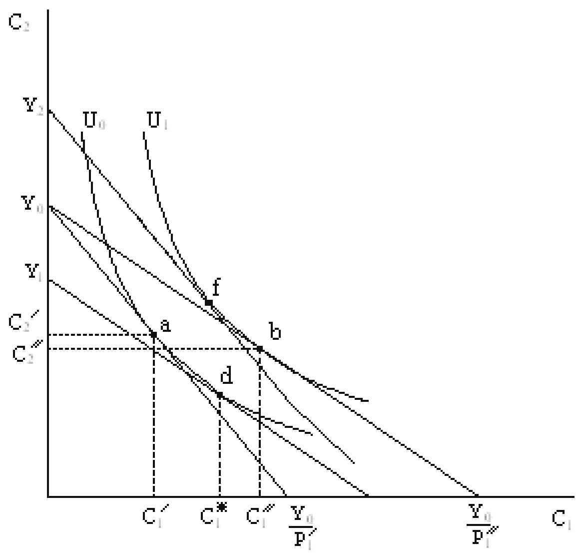
[272] In Figure A.7, C_2 is considered the numeraire good (i.e., prices are adjusted so that P_2' is equal to 1).

[273] For a review of the utility maximizing behavior of consumers, see any general microeconomics textbook.

[274] In Figure A.8, this would result in a shift from C_1'' to C_1^*. This is known as the *income effect* of the price change. The shift from C_1' to C_1^* is considered the *substitution effect*.

CV and EV are simply measures of the distance between the two indifference curves. However, the amount of money associated with CV, EV, and Marshallian consumer surplus (MCS) is generally not the same. For a price fall, it can be shown that CV < MCS < EV, and for a price increase, CV > MCS > EV.[275] Notice that in the case of a price decrease, the CV measures how much the consumer would be willing to pay (WTP) to receive the price reduction and EV measures how much the consumer would be willing to accept (WTA) to forgo the lower price. If the price of C_1 were to increase, then the relationships between WTP/WTA and CV/EV would be reversed. CV would measure the consumer's WTA to suffer the price increase and EV would be the individual's WTP to avoid the increase in price.

Figure A.8



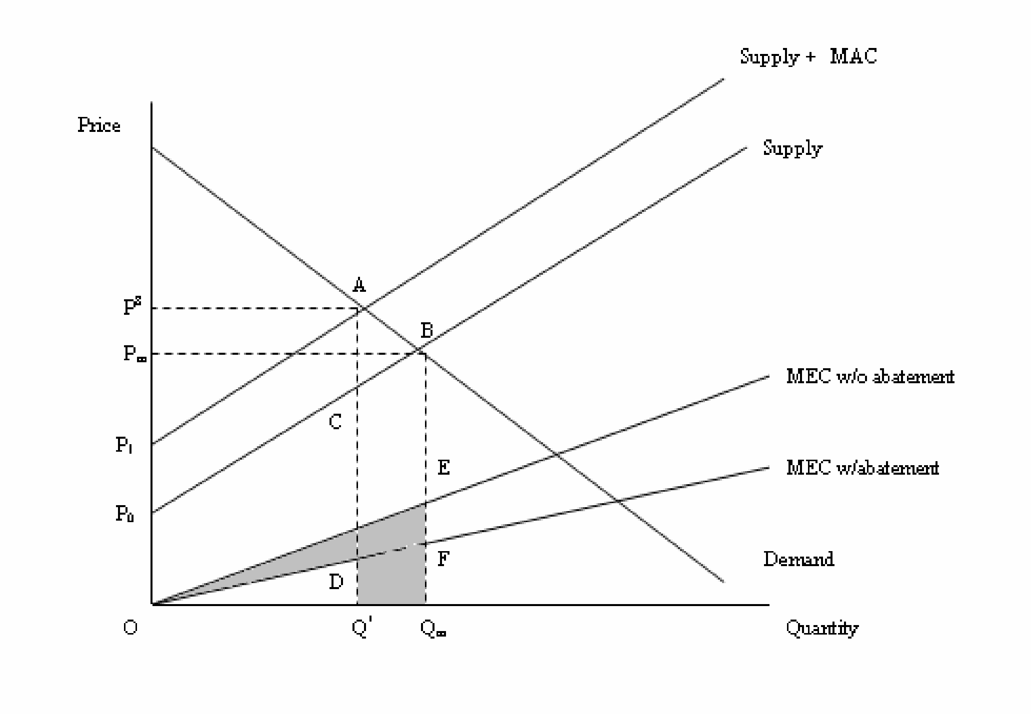
In order to examine the implications of these measures for valuing changes in environmental conditions, one can think of C_1 in the above discussion as an environmental commodity, henceforth denoted by *E*. Then an improvement in environmental quality (or an increase in an environmental public good) resulting from some policy is reflected by an increase in the amount of *E*. Holding all else constant, such an increase is equivalent to a decrease in the price of *E* and can be depicted as a shifting outward of the budget line along the *E* axis.

Welfare changes due to an increase in E follow along the lines of the previous discussion. However, because *E* is generally non-exclusive and non-divisible, the consumer consumption level cannot be adjusted. Therefore, the associated monetary measures of the welfare change are not technically CV and EV, but are referred to as *compensating surplus* (CS) and *equivalent surplus* (ES). In practice, however, the process is the same; a Hicksian demand curve is estimated for the unpriced environmental good. Analogous to the preceding discussion, if there is an environmental improvement, then CS measures the amount of money the consumer would be willing to pay (WTP) for the improvement that would result in the pre-improvement level of utility. For the purposes of environmental valuation, this is the primary measure of concern when considering environmental improvements. ES measures how much society would have to pay the consumer to give him the same utility as if the improvement had occurred. In other words, this is how much he would be willing to accept (WTA) to not experience the gain in

[275] This can be seen by redrawing Figure A.8 using a graph of Marshallian and Hicksian demand curves. See Perman et al. (2003) for a detailed explanation.

environmental quality. If valuing an environmental degradation, then CS measures the WTA and ES measures WTP.

Figure A.9



Whereas statements can be made about the relative size of CV, EV, and MCS for price changes of normal goods,[276] it is not possible to make similar statements about CS, ES, and MCS for a change in environmental quality (Bockstael and McConnell, 1993). Given that environmental quality is generally an unpriced public good, ordinary Marshallian demand functions cannot be estimated, so it may seem irrelevant that one cannot say anything about how MCS approximates the proper measure. However, Bockstael and McConnell's results are important in relation to indirect methods for environmental valuation. However, most indirect valuation studies are based on Marshallian demand functions in practice, in the hope of keeping the associated error small.

A.4.3 Single Market, Multi-Market, and General Equilibrium Analysis

Both supply and demand elasticities are affected by the availability of close complements and substitutes. This highlights the fact that regulating one industry can have an impact on other, non-regulated markets. However, this does not necessarily imply that all of these other markets must be modeled. Changes due to government regulation can be captured using only the equilibrium supply and demand curves for the affected market, assuming (1) there are small, competitive adjustments in all other markets, and (2) there are no distortions in other markets. This is referred to as *partial equilibrium analysis*.

For example, suppose a new environmental regulation increases per unit production costs. The benefits and costs of abatement in a partial equilibrium setting can be illustrated in Figure A.9 where the market

[276] Willig (1976) shows that ordinary, or Marshallian, demand curves may provide an approximate measure of welfare changes resulting from a price change. In most cases, the error associated with using MCS, with respect to CV or EV, will be less than 5% (see Perman et al., 2003).

produces the quantity Q_m in equilibrium without intervention. The external costs of production are shown by the marginal external costs (MEC) curve without any abatement. Total external costs are given by the area under the MEC curve up to the market output, Q_m, or the area of triangle Q_mE0.

With required abatement production, costs are the total of supply plus marginal abatement costs (MAC), shown as the new, higher supply curve in the figure. These higher costs result in a new market equilibrium quantity shown as Q_1. The social cost of the requirement is the resulting change in consumer and supplier surplus, shown here as the total observed abatement costs (parallelogram P_0P_1AC) plus the area of triangle ABC, which can be described as deadweight loss.

Abatement also produces benefits by shifting the MEC curve downward, reflecting the fact that each unit of production now results in less pollution and social costs. Additionally, the reduced quantity of the output good also results in reduced external costs. The reduced external costs, i.e. the benefits, are given by the difference between triangle Q_mE0 and triangle Q^*D0, represented by the shaded area in the figure.

The net benefits of abatement are the benefits (the reduced external costs) minus the costs (the loss in consumer and producer surplus). In the figure this would equal the shaded area (the benefits) minus total abatement costs and deadweight loss as described above.

While the single market analysis is theoretically possible, it is generally impractical for rulemaking. As was mentioned in Section A.3, this is often because the gains occur outside of markets and cannot be linked directly to the output of the regulated market. Therefore, BCA is frequently done as two separate analyses: a benefits analysis and a cost analysis.

When a regulation is expected to have a large impact outside of the regulated market, then the analysis should be extended beyond that market. If the effects are significant but not anticipated to be widespread, one potential improvement is to use multi-market modeling in which vertically or horizontally integrated markets are incorporated into the analysis. The analysis begins with the relationship of input markets to output markets. A multi-market analysis extends the partial equilibrium analysis to measuring the losses in other related markets.[277]

In some cases, a regulation may have such a significant impact on the economy that a general equilibrium modeling framework is required.[278] This may be because regulation in one industry has broad indirect effects on other sectors, households may alter their consumption patterns when they encounter increases in the price of a regulated good, or there may be interaction effects between the new regulation and pre-existing distortions, such as taxes on labor. In these cases, partial equilibrium analyses are likely to result in an inaccurate estimation of total social costs. Using a general equilibrium framework accounts for

[277] An example of the use of multi-market model for environmental policy analysis is contained in a report prepared for EPA on the regulatory impact of control on asbestos and asbestos products (EPA, 1989).

[278] *General equilibrium analysis* is built around the assumption that, for some discrete period of time, an economy can be characterized by a set of equilibrium conditions in which supply equals demand in all markets. When this equilibrium is "shocked" through a change in policy or a change in some exogenous variable, prices and quantities adjust until a new equilibrium is reached. The prices and quantities from the post-shock equilibrium can then be compared with their pre-shock values to determine the expected impacts of the policy or change in exogenous variables.

linkages between all sectors of the economy and all feedback effects, and can measure total costs comprehensively.[279]

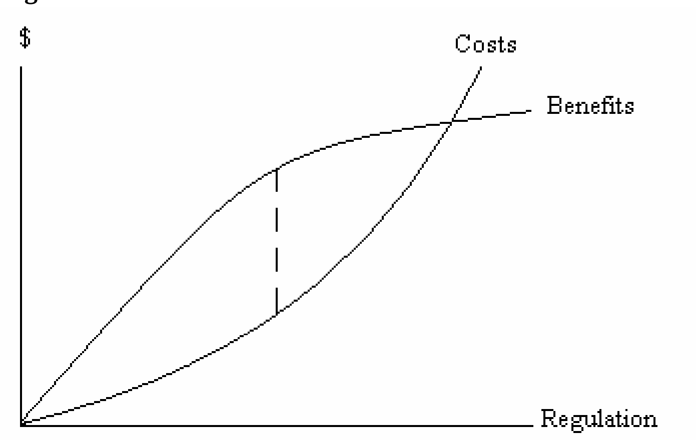
A.5 Optimal Level of Regulation

Following from the definition in Section A.1, the most economically efficient policy is the one that allows for society to derive the largest possible social benefit at the lowest social cost. This occurs when the *net* benefits to society (i.e., total benefits minus total costs) are maximized. In Figure A.10, this is at the point where the distance between the benefits curve and the costs curve is the largest and positive.

Note that this is *not* necessarily the point at which:

- Benefits are maximized,
- Costs are minimized,
- Total benefits = total costs (i.e., benefits/costs ratio = 1),
- Benefits/costs ratio is the largest, or
- The policy is most cost-effective.

Figure A.10



If the regulation were designed to maximize benefits, then any policy, no matter how expensive, would be justified if it produced any benefit, no matter how small. Similarly, minimizing costs would, in most cases, simply justify no action at all. A benefits/costs ratio equal to one is equivalent to saying that the benefits to society would be exactly offset by the cost of implementing the policy. This implies that society is indifferent between no regulation and being regulated; hence, there would be no net benefit from adopting the policy. Maximizing the benefits/costs ratio is not optimal either. Two policy options could yield equivalent benefits/costs ratios but have vastly different net benefits. For example, a policy that cost $100 million per year but produced $200 million in benefits has the same benefit/cost ratio as a policy that cost $100,000 but produced $200,000 in benefits, even though the first policy produces substantially more net benefit for society.[280] Finally, finding the most cost-effective policy has similar

[279] Chapter 8 provides a more detailed discussion of partial equilibrium, multi-market, and general equilibrium analysis.

[280] However, benefit-cost ratios are useful when choosing one or more policy options subject to a budget constraint. For example, consider a case where five options are available and the budget is $1,000. The first option will cost $1,000 and will deliver benefits of $2,000. Each of the other four will cost $250 and deliver benefits of

problems because the cost-effectiveness ratio can be seen as the inverse of the benefit/cost ratio. A policy is cost effective if it meets a given goal at least cost – i.e., minimizes the cost per unit of benefit achieved. Cost effectiveness analysis (CEA) can provide useful information to supplement existing BCA and may be appropriate to rank policy options when the benefits are fixed and cannot be monetized, but it provides no guidance in setting an environmental standard or goal.

Conceptually, net social benefits will be maximized if regulation is set such that emissions are reduced up to the point where the benefit of abating one more unit of pollution (i.e., marginal social benefit)[281] is equal to the cost of abating an additional unit (i.e., marginal abatement cost).[282] If the marginal benefits are greater than the marginal costs, then additional reductions in pollution will offer greater benefits than costs, and society will be better off. If the marginal benefits are less than marginal costs, then additional reductions in pollution will cost society more than it provides in benefits, and will make it worse off. When the marginal cost of abatement is equal to society's marginal benefit, no gains can be made from changing the level of pollution reduction, and an efficient aggregate level of emissions is achieved. In other words, *a pollution reduction policy is at its optimal, most economically efficient point when the marginal benefits equal the marginal costs of the rule.*[283]

The condition that marginal benefits must equal marginal costs assumes that the initial pollution reduction produces the largest benefits for the lowest costs. As pollution reduction is increased (i.e., regulatory stringency is increased), the additional benefits decline and the additional costs rise. While it is not always true, a case can be made that the benefits of pollution reduction follow this behavior. The behavior of total abatement costs, however, will depend on how the pollution reduction is distributed

---

$750. If options are selected according to the net benefits criterion, the first option would be selected, because its net benefits are $1,000 while the net benefits of each of the other options are $500. However, if options are selected by the benefit-cost ratio criterion, the other four options would be selected, as each of their benefit cost ratios equal 3, versus a benefit-cost ratio of 2 for the first option. In this case, choosing options by the net benefits criterion would yield $1,000 in total net benefits, while choosing options by the benefit-cost ratio criterion would yield $500 in total net benefits. In most cases, choosing options in decreasing order of benefit-cost ratios will yield the largest possible net benefits given a fixed budget. (This method will guarantee the optimal solution if the benefits and costs of each option are independent, and if each option can be infinitely subdivided: simply select the options in decreasing order of their benefit-cost ratios and once the budget is exceeded subdivide the last option selected such that the budget constraint is met exactly (e.g., see Dantzig (1957)).) Also note that this strategy does not require measuring benefits and costs in the same units, which means that it is directly useful for cost-effectiveness analysis (e.g., Hyman and Leibowitz (2000)), while the net-benefit criterion is not.

[281] The benefits of pollution reduction are the reduced damages from being exposed to pollution. Therefore, the marginal social benefit of abatement is measured as the additional reduction in damages from abating one more unit of pollution.

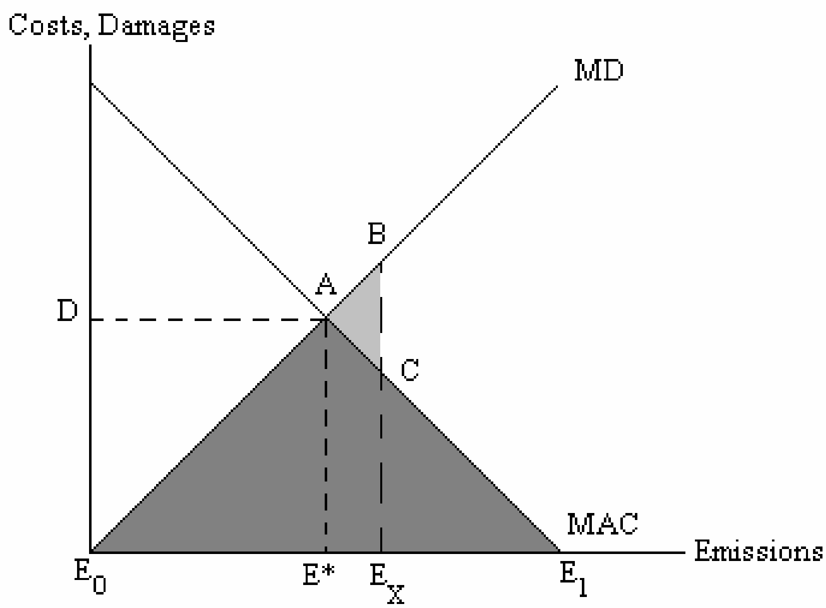
[282] The idea that a given level of abatement is efficient – as opposed to abating until pollution is equal to zero – is based on the economic concept of diminishing returns. For each additional unit of abatement, marginal social benefits decrease while marginal social costs of that abatement increase. Thus, it only makes sense to continue to increase abatement until the point where marginal abatement benefits and marginal costs are just equal. Any abatement beyond that point will incur more additional costs than benefits. (Alternatively, one can understand the efficient level of abatement as the amount of regulation that achieves the efficient level of pollution. If one considers a market for pollution, the socially efficient outcome would be the point where the marginal willingness to pay for pollution equals the marginal social costs of polluting.)

[283] It is important to reemphasize the word "marginal" in this statement. Marginal, in economic parlance, means the extra or next unit of the item being measured. If regulatory options could be ranked in order of regulatory stringency, then marginal benefits equal to marginal costs means that the additional benefits of increasing the regulation to the next degree of stringency is equal to the additional cost of that change.

among the polluters since firms may differ in their ability to reduce emissions. The aggregate marginal abatement cost function shows the least costly way of achieving reductions in emissions. It is equal to the horizontal sum of the marginal abatement cost curves for the individual polluters. Although each firm faces increasing costs of abatement, marginal cost functions still vary across sources. Some firms may abate pollution relatively cheaply, while others require great expense. To achieve economic efficiency, the lowest marginal cost of abatement must be achieved first, and then the next lowest. Pollution reduction is achieved at lowest cost only if firms are required to make equiproportionate cutbacks in emissions. That is, at the optimal level of regulation, the cost of abating one more unit of pollution is equal across all polluters.[284]

Figure A.11 illustrates why the level of pollution that sets the marginal benefits and marginal costs of abatement equal to each other is efficient. [285] Emissions are drawn on the horizontal axis and increase from left to right.. The damages from emissions are represented by the marginal damage curve (MD). Damages may include the costs of worsened human health, reduced visibility, lower property values, and loss of crop yields or biodiversity. As emissions rise, the marginal damages increase. E_1 represents the amount of emissions in the absence of regulation on firms. The costs of controlling emissions are represented by the marginal abatement cost curve (MAC). As emissions are reduced below E_1, the marginal cost of abatement rises.

Figure A.11



[284] Thus a regulation that requires all firms to achieve the same level of reduction will probably result in different marginal costs for each firm and not be efficient. (See Field and Field (2002), p. 105, or any other environmental economics text for a detailed explanation and example.)

[285] Figure A.11 illustrates the simplest possible case, where the pollutant is a flow (i.e., it does not accumulate over time) and marginal damages are independent of location. When pollution levels and damages vary by location, then the efficient level of pollution is reached when marginal abatement costs adjusted by individual transfer coefficients are equal across all polluters. Temporal variability also implies an adjustment to this equilibrium condition. In the case of a stock pollutant, marginal abatement costs are equal across the discounted sum of damages from today's emissions in all future time periods. In the case of a flow pollutant, this condition should be adjusted to reflect seasonal or daily variations (see Sterner (2003)).

The total damages associated with emissions level E^* are represented by the area of the triangle AE_0E^*, while the total abatement costs are represented by area AE_1E^*. The total burden on society of this level is equal to the total abatement costs of reducing emissions from E_1 to E^* plus the total damages of the remaining emissions, E^*. That is, the total burden is the darkly shaded triangle, E_0AE_1.

Now assume that emissions are something other than E^*. For example, suppose emissions were E_X, which is greater than E^*. In this case, total damages for this level of emissions are equal to the area of the triangle BE_0E_x, while total costs of abatement to this level is equal to the area CE_xE_1. The total burden on society of this level is the sum of the areas of the darkly shaded and the lightly shaded triangles. This means that the excess social cost of choosing emissions E_X rather than E^* is equal to the area of the lightly shaded triangle, ABC. A similar analysis could be done if emissions levels were below level, E^*. Here, the additional abatement costs would be greater than the decrease in damages, resulting in excess social costs. The policy that sets the emissions level at E^* – at the point where marginal benefits of pollution reduction (represented by the marginal damage (MD) curve) and the marginal abatement cost (MAC) curve intersect – is economically efficient because it imposes the least net cost on (i.e., yields the highest net benefits for) society. That is, the triangle E_0AE_1 is the smallest shaded region that can be obtained.

A.6 Conclusion

The purpose of this appendix is to present a brief explanation of some of the fundamental economics relevant to Chapters 3 through 9. It is not intended to provide a comprehensive discussion of all microeconomic theory and its application to environmental issues. The interested reader can turn to undergraduate or graduate level textbooks for a more thorough exposition of the topics covered here. At the undergraduate level, Field and Field (2002) provide an introduction to the basic principles of environmental economics. Tietenberg's (2002) and Perman et al.'s (2003) presentations are more technical but still used primarily for undergraduate courses. Freeman (2003) is the standard text for graduate courses in environmental economics and deals with the methodology of non-market valuation. Supplemental texts that provide a good handle on environmental economics with less technical detail include Stavins (2000) and Portney and Stavins (2000). Finally, general microeconomics textbooks (Mankiw (2004), Varian (2005) at the undergraduate level, and Mas-Colell et al. (1995), Kreps (1990), Varian (1992) at the graduate level), and applied welfare economics textbooks (Just et al., 2005) are useful references as well.