> The man who chooses to work longer to gain an income more than sufficient for his basic needs prefers some extra goods or services to the leisure and activities he could perform during the possible nonworking hours; whereas the man who chooses not to work the extra time prefers the leisure activities to the extra goods or services he could acquire by working more.
>
> **Given this, if it would be illegitimate for a tax system to seize some of a man's leisure (forced labor) for the purpose of serving the needy , how can it be legitimate for a tax system to seize some of a man's goods for that purpose?** Why should we treat the man whose happiness requires certain material goods or services differently from the man whose preferences and desires make such goods unnecessary for his happiness? [Nozick, Robert (1974 / 2013). *Anarchy, State, and Utopia* (p. 170).]

I. Introduction: Neoclassical Economics and Normative Analysis in the Twentieth Century

During the late nineteenth century and early twentieth century, a series of innovations in economics generated new geometric and mathematical representations of rational persons, their choices, individual markets, and market systems. The models included utility functions, indifference curve representations of consumers, profit maximizing models of a firm's output decisions, demand and supply curve representations of individual market prices, and general equilibrium models of an economy's prices. These innovations deepened our understanding of how market systems operated.

These same tools could and were used to deepen our understanding of the normative properties of commerce and commercial systems. Utilitarians and welfare economists could use the models developed to characterize aggregate utility or social net benefits for specific types of markets and market settings. They could also use them to assess the relative merits of alternative public policies with respect to markets, which allowed policy recommendations to be made that were systematically grounded in these normative theories and economic analysis.

By mid century, a standard neoclassical "tool bag" of geometric and mathematical models for microeconomics and welfare economics had been assembled and was routinely taught at most universities.[1] The utility concept itself came to play a fundamental role in economics. The diminishing marginal utility concept was used to explain the diamond-water paradox. The idea the persons would maximize utility was used to model the thought processes that produce consumer demand, firm owner decisions, market prices, and personal income. Mathematical and geometric representations of utility were used to illustrate the existence of gains to trade. The new mathematical representations of aggregate utility also allowed the effects of economic development and public policy on aggregate utility to be rigorously deduced from economic models.

Contractarian assessments of markets and institutions using neoclassical tools arose later, but were also facilitated by them. During the 1950s and 1960s Rawls and Buchanan routinely use neoclassical concepts and geometry to assess the relative merits of alternative legal and political institutions. There was no necessary conflict between the interests of an individual and pursuit of the good society. It was not necessary to add up mental states to demonstrate that everyone can benefit from institutions that address social dilemma or promote voluntary exchange. The various Pareto principles were often used to formalized this non-utilitarian intuition. One state of the world (A) is better than (Pareto superior to) another (B), if at least one person prefers A to B and no one prefers B to A.

Utility-based, as opposed to money-based representations of exchange clearly demonstrate that each participant prefers the result after the exchange to that before the exchange. If no one else is harmed by the

[1] The precision of the new geometric and mathematical analyses exaggerates their ability to shed light on economic and social activities, but may important insights were developed using the neoclassical approach. Indeed, the precision associated with the new analysis caused the older more philosophical and historical approaches of nineteenth century economics to gradually disappeared from mainstream economics.

transaction, the result can be regarded as an improvement, a better society. This is the normal consequence of any enterprise adopted through contract and unanimous agreement. Nonetheless, Pigou's money-based applied utilitarianism dominated normative analysis for most of the twentieth century.

Although specialization in academia radically increased during the twentieth century, there remained an overlap between the work of economists and philosophers on issues of system choice and the good society. Many of the leading utilitarians of the nineteenth century were also highly regarded for their contributions to economics. Bentham, Mill, and Sidgwick, for example, all made significant contributions to economics through their political economy books. Mill's textbook was among the most used in the English speaking world during the mid nineteenth century. Among the twentieth century philosophers, Rawls, Nozick, and Skyrms, for example, are well known for their use of rational-choice based analysis to develop conclusions about the nature of the good society. Among well-known twentieth century economists, Buchanan, Harsanyi, and Sen are, for example, also known for their many contributions to normative theory.

This chapter show how the rational-choice approach can be used assess the relative merits of markets and public policies. The chapter analyzes both strengths and weaknesses of market systems from the vantage point of welfare economics. The models applied are somewhat more demanding than those used in Part II. The appendix to chapter 10 provides a short review of the assumptions, logic, geometry, and mathematics behind the models used.

II. Utility, Indifference Curves, and Mutual Gains to Trade

Economists usually take a materialistic view of utility rather than a philosophical one. The consumption of goods and services is the ultimate source of utility in economic models, rather than ethical dispositions as in Aristotle and Smith. Given this materialistic assessment of the interests of most consumers and the usual assumptions about preferences (complete and transitive), it is possible to characterize utility as a function from combinations of goods and services into real numbers that represent utility levels, as with U= u(X, Y, Z, ...).

Economists also make assumptions about the manner in which utility is generated by goods and services. For example, it is assumed that goods exhibit diminishing marginal utility: each successive unit of a good generates less and less additional utility. This implies that marginal utility curves are downward sloping. If the consumption of one good does not affect the marginal utility of any other (separability of function u), this assumption is sufficient to prove that an individual's demand curves also slope downward, and that the utility function itself is strictly concave, both important mathematical ideas in economics.[2]

The mathematics of utility functions allows choices over any finite number of goods to be represented as a constrained maximization problem. Typically, consumers are assumed to maximize utility subject to budget constraint. Fortunately, two-dimensional choice settings are sufficient to illustrate the main implications of the decisions made by utility maximizing persons, what economists usually mean by rational persons. (The term rationality has quite different meanings for economists, philosophers, and psychologists.)

This chapter uses geometric representations of rational choice and markets to analyze markets and public policies. Utility functions can be graphed in the same way that any other function can. For purposes of illustration, however, it is often convenient to reduce the number of goods to two and graph the three dimensions (utility, good X, and good Y) in goods space along by drawing the contour lines of the function. These contour lines came to be called indifference curves, because they represent different combinations of goods that generate the same utility levels.

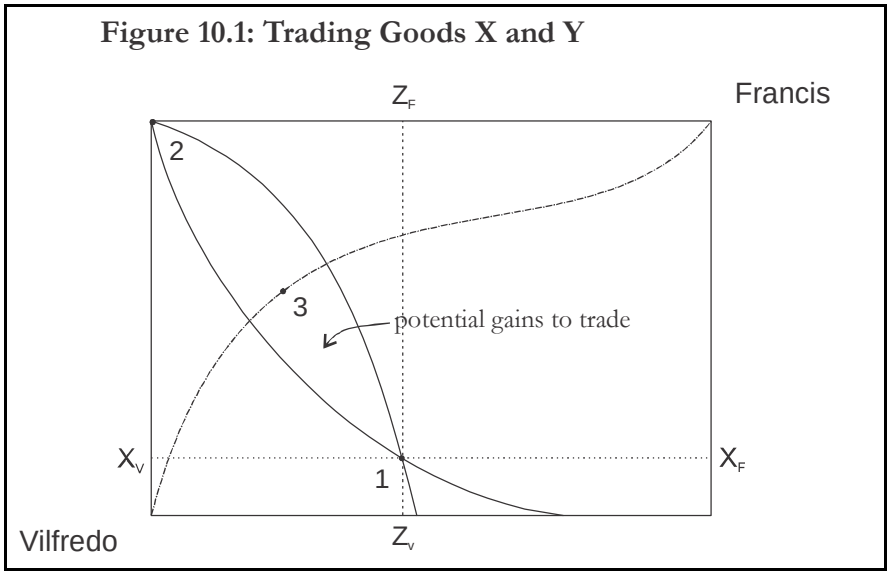
[2] Some forms of interdependency among goods are also consistent with downward sloping demand curves and concavity of utility functions. If the consumption level of one good increases the marginal utility of other goods, demand curves will also always slope downward. Strict concavity implies that a function can have at most one maximum. Utility functions that are strictly concave, continuous, and differentiable have indifference curves that are "c-shaped" or "o-shaped," as drawn in most economic text books and in figure 10.1.

A. Indifference Curves and Gains to Trade

A two-good utility function U = u(X,Z) can be represented in a two dimensional diagram by graphing all the combinations of X and Z that generate a particular utility level. Each such line or curve is called an indifference curve, becasue rational decision makers are indifferent between any two possibilities that generate the same utility.. The "higher" an indifference curve a person can reach, the higher is their utility, and therefore the greater is their welfare or happiness.

In 1881, an economists named Francis Edgeworth realized that one could represent gains to trade using indifference curves. In 1906 Vilfredo Pareto realized that one could represent trade between two persons by using indifference curves to represent a choice setting in which given amounts of two goods are held by two persons. The result was called an Edgeworth box. The Edgeworth box very nicely illustrates the idea of mutual gains to trade from both utilitarian and contractarian perspectives.

Figure 10.1: Trading Goods X and Y



Suppose that finite quantities of two goods, X and Z, are initially distributed between Vilfredo and Francis. For example, at position 1, Francis and Vilfredo each have positive quantities of both goods. Since both prefer more to less of each good, it might appear their preferences necessarily place them in conflict with each other; however, this is not necessarily the case. In the diagram, Francis becomes better off (reaches higher indifference curves and therefore utility levels) as divisions of the good move to the south west and Vilfredo become better off as divisions of the two goods move to the northeast. It turns out that for most preferences and divisions of two goods, there are regions inside the Edgeworth where both parties can simultaneously be made better off. The lens shaped region between the indifference curves passing through division 1 is such a region.

A move from to any point in the lens-shaped area labeled "gains to trade" will simultaneously place both Vilfredo and Francis on a higher indifference curve. These points can be reached if Vilfredo trades some of his good Z for some of Francis' good X. Within such regions, mutual gains to trade can be realized by Vilfredo and Francis, even in a "zero-sum" world without production. Voluntary exchange, even in a barter based society, makes everyone better off.

There are a few implicit assumptions in characterizing the Edgeworth box. For example, it is assumed that initial ownership rights include the right to transfer goods, and that the law is sufficiently well enforced that theft or conquest are ruled out. Given those institutional assumptions, self interest is sufficient to induce increases in aggregate utility in such cases.

B. Normative Implications of Gains to Trade

Gains to trade are exhausted when divisions of goods reach any point on the wavy line running from Francis' corner to Vilfredo's corner (the contract curve). Along that line, the indifference curves of Francis and Vilfredo are tangent to one another. This implies that there is no lens-shaped area of gains to trade associated with such points. Points on the contract curve are Pareto optimal, no shift in resources are possible that can make one person better off without making another worse off.

The Edgeworth box can also be used to illustrate differences between the contractarian and utilitarian principles. The fact that higher indifference curves are reached through trade clearly implies that aggregate utility is increased by trade. Trade is thus good from the perspective of utilitarianism. However, the possibility of such gains also implies that aggregate utility is not maximized at all possible divisions of resources.

Utilitarians can thus recommend a wide variety of alternative divisions of the resources that would increase aggregate utility.

Without precise knowledge of the individual utility functions, most such divisions would fail to maximize aggregate utility and, were the recommendations to be adopted, voluntary exchange take place after such allocations were made. An exception to this rule occurs when Vilfredo and Francis have identical utility functions (a convenient, but unrealistic assumption) and both functions exhibits diminishing marginal utility. In that case, the equal division that lies at the exact center of the box will both be on the contract curve and maximize aggregate utility. Except in that one case, precise measures of utility are necessary to determine the best utilitarian allocation of the resources.

Contractarian analysis assumes that a legitimate sphere of individual actions (rights) exist prior to a social contract. In most cases, these including the rights associated with holding property (including oneself) and the right to transfer some of one's rights to another. The latter is used to create a social contract under most contractarian theories. Commerce also relies on this ability. Voluntary transfers of rights from one person to another advance the common interests of those engaging in exchange. Voluntary exchange thus tends to satisfy contractarian norms. Voluntary exchange necessarily increases utility, whereas mandating shifts of goods from one person to the other is not likely to do so unless a good deal is known about individual preferences and the aim is to increase aggregate utility.

C. Gains to Trade and Property Institutions

The Edgeworth box also illustrates why ownership rights and the right to transfer those rights tends to be part of both contractarian and utilitarian legal systems. For contractarians, institutions are chose from behind an veil of ignorance or uncertainty. In such cases, no one knows what their position will be in the society that emerges until after the institutions are chosen. Institutions that allow voluntary exchange to take place allow everyone to trade goods they own (including their own labor) for others they do not in a manner that makes both parties better off. Since everyone anticipates advantages from such rights, they would be agreed to. Exceptions might exist for some goods (as with slavery or contracts for murder), but in general institutions that support trade tend to benefit everyone.[3]

For utilitarians, institutions that enable trade to take place are also advantageous. Trade tends to increase aggregate utility, because all the parties to voluntary exchange benefit. Such rights would arguably be unnecessary if all decisions were made by utilitarian social planners with perfect information, but no utilitarian (as far as this author knows) believes that such an institution is actually possible. Failing that ideal, gains to trade are likely to be commonplace and voluntary exchange provides a mechanism through which those gains can be realized and aggregate utility increased.

In sum, commerce will play a role in every society that is grounded on contractarian or utilitarian principles.

III. Welfare Implications of the Net Benefit Maximizing Model

An alternative to the utility maximizing model of rationality is the net benefit maximizing model. The net-benefit model of rational choice can be used to characterize choices of consumers and firms, supply and demand, and market equilibrium. (The essential mathematics of net benefit maximizing choices are developed in the appendix.)

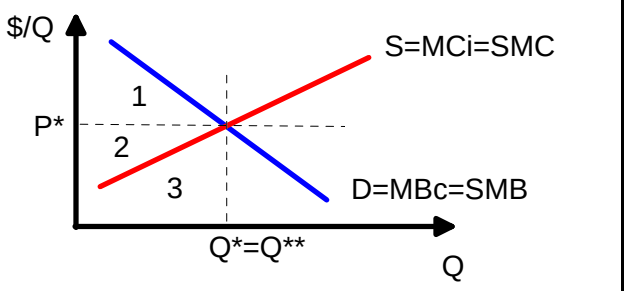
A person's net benefits from market exchange can be used as an index of his or her utility. These, being in dollar terms can be added up to create social net benefits, which is dollar-based measure of aggregate utility. Much of applied welfare economics attempt to assess whether markets maximize social net benefits and, if not, whether policies exist that can increase social net benefits.

[3] This conclusion ignores the possibility of internalized norms that are anti-trade or strongly egalitarian. It also ignores problems associated with malice and envy. However, even in such cases, some trades make the parties involved better off, and rules that enable such trades would be favored. (Even in Thomas More's anti-commercial utopia, trade took place between utopia and outsiders.) Externality problems issues are also ignored in the above discussion; however, these tend affect the nature of ownership rights (use rights) rather than the transferability of such rights.

There are, of course, differences between a net benefit representation and a utility representation of both rational choice and the outcomes of actions taken. As Pigou noted in 1920:

> [T]he money which a person is prepared to offer for a thing measures directly, not the satisfaction he will get from the thing, but the intensity of his desire for it. This distinction, obvious when stated, has been somewhat obscured for English speaking students by the employment of the term utility —which naturally carries an association with satisfaction— to represent intensity of desire.
>
> Thus, when one thing is desired by a person more keenly than another, it is said to possess a greater utility to that person. (*The Economics of Welfare*, KL 483-486).

Figure 10.2 Market Equilibrium



It is a great convenience for elementary welfare economics that the social marginal benefits of consumption can often be represented with market demand curves and the marginal social cost of production with long run market supply curves, under conditions that are not much stronger than those required to derive them in the first place, as developed in the appendix.

The areas between the demand and supply curves correspond to social net benefits, as long as all the benefits from this product accrue to consumers in the market of interest, and production costs are fully accounted for by each firm's own cost functions. In that case, it can be easily shown that competitive markets tend to produce output levels that maximize social net benefits.

A. Demand, Supply, and the Normative Merits of Competitive Markets

Figure 10.2 illustrates a typical competitive market equilibrium. In a market without externalities, all the relevant costs and benefits for a single market are characterized by the demand and supply curves. The total benefits realized by consumers can be approximated as the area under the demand curve out to the quantity of interest (here the sum of the areas 1+2+3).[4] The cost of the product to consumers is the area under the price line (areas 1+2), and the net benefits realized by consumers (consumer surplus) is thus area 1, the difference between aggregate consumer benefits and their costs.

The profits realized by firms is also a net benefit. The industry's total revenue is the area under the price line from zero to the quantity of interest (here areas 1+2) and industry's total cost can be approximated as the area under the supply curve from zero to the quantity of interest (here area 3). Industry profits are the difference between revenues and costs, which is area 2 in the diagram. Social net benefits are the sum of consumer surplus (area 1) and profits (area 2). Note that in the absence of externalities, social net benefits are areas between the demand and supply curves.[5]

[4] When demand curves are derived from utility functions, the demand curves should be of the Hicksian compensated demand variety. However, it is possible to directly derive demand curves from individual consumer marginal benefit curves. When that method is used (which is illustrated in the appendix), individual demand curves and marginal benefit curves go through exactly the same points and thus such demand curves can be used as marginal benefit curves.

[5] Ricardian rather than Marshallian long run supply curves are used in the diagrams of this and the next chapter, partly for geometric convenience, but also because they characterize a broad range of markets. In the Ricardian case, industry long run supply can be interpreted as the industry's marginal cost

The purpose of production from a utilitarian perspective is arguably the benefits that consumers gain from it.[6] In this case, which is the usual assumption, the social marginal benefit curve can be approximated with the market demand curve (in the absence of externalities) and the social marginal cost curve can be approximated with the supply curve. As true of other efforts to maximize net benefits, maximizing social net benefits requires an output level, Q**, where social marginal benefits equal social marginal costs. Since this occurs where supply crosses demand, the output normally produced by competitive markets tends to maximize social net benefits.

Subject to various caveats, Figure 10.3 thus demonstrates that competitive markets tend to maximize aggregate welfare measured as represented with social net benefits.[7] Insofar as social net benefits are correlated with aggregate utility, competitive markets also tend to maximize aggregate utility for a given distribution of income.[8]

Market prices determine the division of social net benefits. The competitive equilibrium divides social net benefits between firm owners (area 2) and consumers (area 3) is determined by equilibrium price and the slopes the relevant demand and supply curves. The sum of the net benefits realized by consumers and firms is another meaning of the term "social dividend," which differs from the usual one that focuses on the market value of the final goods and services produced and sold.

Figure 10.2 can also be used as a basis for contractarian analysis of competitive markets. The diagram implies that the gains in trade that exist in the barter setting of an Edgeworth box also exist in more developed markets, with many buyers and sellers of goods and services benefiting. As in the Edgeworth Box, all persons can expect to benefit from such market activities in the society that emerges after institutions are chosen. Insofar as essential all persons are buyers and sellers of goods and services, there would be essentially universal support for the institutions that enable competitive markets. The extent of the net benefits realized and of the market that emerges would partly depend on the internalized norms of the individuals in the community, as developed in Part II.

---

function. This contrasts with the Marshallian model, where long run supply can be regarded as an industry's long run average cost function. Ricardian models assume that firms producing the same products have different cost functions. Marshallian models assume that all firms in the industry have the same cost functions.

In general, production cost differences often emerge because of unique features of economic organizations: staff, location, and access to natural resources of varying quality, as in the oil, mining, and agricultural industries. Long run competition does not completely eliminate profits in Ricardian industries. Some firms earn more profits than others.

[6] Economists normally assume that production is undertaken only because of the money rewards of selling the goods and service produced. However, some forms of production can be valuable in themselves. They may, for example, contribute to character development, be entertaining for the producers themselves, yield praise, status, or fame. Taking account of such benefits--character development, entertainment, praise etc--would alter the net-benefits of production and the quantities produced at various prices, as illustrated in chapter 7. Such benefits would also alter the social net benefits of alternative output levels, increasing them in cases in which productive activities are deemed virtuous or otherwise increase happiness, and reducing them if deemed immoral or otherwise diminishing happiness.

[7] The Pareto efficiency of perfectly competitive market outcomes is easy to demonstrate mathematically. See Varian (1992) for a relatively clear illustrating proof. Figure 10.3 can also be used to demonstrate the fundamental theorem of welfare economics. Net gains to trade exist for all units produced having a marginal benefit for consumers that is greater than its marginal production cost. Such potential gains exist for all units up to Q*, and for none afterwards. At the market equilibrium, all potential gains to trade are realized and no changes in output or price are possible that would make one person better off without making another worse off.

[8] The social net benefit approach neglects distributional issues, because it uses dollars rather than utils as the common indicator of happiness. A millionaire's last dollar of net benefits is treated just like a pauper's. Distributional issues are analyzed later in the chapter.

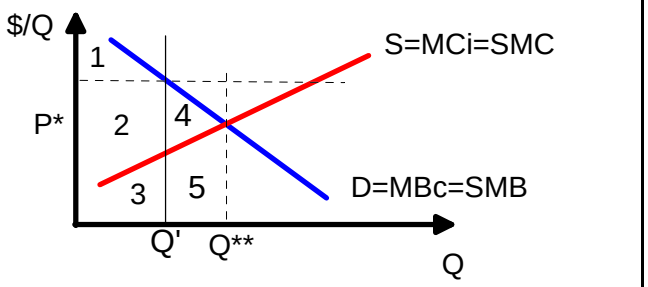
If the material comforts and amusements produced by markets are reinforced by norms that emphasize team work, voluntariness, or entrepreneurship, the economic system favored would have essentially unfettered commerce. It must be acknowledged, however, that in a community composed of individuals whose norms favor non-material goals over material ones, it is possible that essentially unanimous agreement to shift toward a society like that envisioned in More's utopia would exist, in spite of the gains from trade that would be lost by doing so.

B. The Welfare Economics Case Against Trade Barriers

Competitive markets have attractive properties from the perspective of welfare economics and utilitarian and contractarian perspectives. One policy implication of this result is that policy makers should not adopt polices that reduce competition, as with policies that limit participation in market or which block it entirely, unless other ethical ends are more important to citizens than the reduction in net benefits generated by such policies.

The simplest way to limit competition is to simply limit supply by reducing allowed outputs of firms in an industry or making it difficult for new firms to enter the market. Such policies tend to reduce the supply of goods and services to levels below those that would have occurred with more open markets. Since production costs are not affected by such policies, potential supply is not directly affected by such policies. Figure 10.3 illustrates the effects of a mandated reduction in the amount that can be brought to market by each potential supplier, a quota system.

Figure 10.3 Market Equilibrium with Quotas



Prices rise to bring demand down to the level of the allowed market supply, now Q'. At this price, which is higher than the competitive price, consumer surplus falls and profits may rise relative to those in competitive markets. Whether profits rise or not depends on the slopes of the demand and supply curves and the nature of the regulations. As drawn, profits rise relative to those in Figure 10.2. In either case, social surplus falls by the area of triangle 4.

Welfare economics thus implies that policies that directly or indirectly reduce amounts that potential suppliers can bring to market are dominated by ones that do not. That is to say, welfare economics implies that policies that create barriers to entry are "bad" policies, other things being equal. This is not to say that other things are always equal, licensing for example could raise the average quality of services, but unless a barrier to trade generates social net benefits greater than area 4, such policies reduce social net benefits and should be avoided.

Insofar as social net benefits are a good proxy for aggregate utility in the markets of interest utilitarians would oppose such barriers--as most utilitarians did in the nineteenth century. Contractarians would also tend to oppose such barriers, but not because of effects on social net benefits but because of expected effects on and individual's gains from trade in the markets of interest. Insofar as persons expect to be worse of in markets with barriers to entry than ones without them, they would prefer more open markets to relatively closed ones.

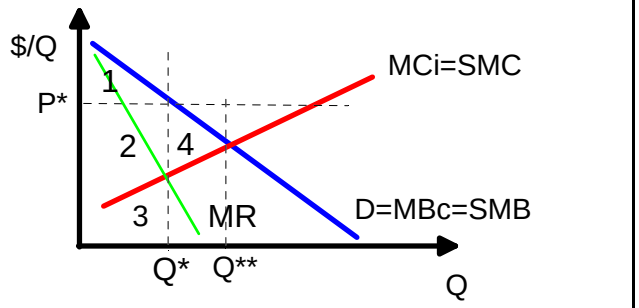
C. Welfare Economics and the Neoclassical Case Against Monopoly

An extreme case of a barrier to entry is the formal grant of a monopoly privilege. In such cases, a single firm is "given" (or purchases) the exclusive right to sell in a particular product in a particular market. The creation of local monopolies is common historically, as for example British kings and queens used to sell monopoly privileges and use the proceeds as royal revenues. This practice declined around 1600, but was still occasionally used. For example, during the twentieth century, state and local governments in the United States often created local monopolies in electricity

supply or distribution, or in other cable-based services. Other monopoly markets may emerge through cartel agreements or because a market is too small to support more than one efficiently sized firm. A firm that has an exclusive right to sell a particular product in a particular market will naturally choose an output with market price, revenues, and profits in mind.

There are two textbook forms of monopoly. First there are monopolists that sell all of their output at a single price. Second, there are monopolists that attempt to extract all the benefits possible from their production by selling its output at different prices to different persons. Both can be analyzed with figure 10.4.

Figure 10.4 Monopoly Equilibrium



In the case in which a monopolist a monopolist sells all of its output at a single price, it will take account of the effect that price has on sales, revenues, and costs and select the profit maximizing price and output combination. The profit maximizing output is found where marginal revenue (MR) equals marginal cost (MCi). Marginal revenue is not price in monopolistic markets, because its selling price necessarily falls with output. The profit-maximizing output is labeled Q* in the diagram. The profit maximizing price is the one that induces consumers to purchase exactly Q* units of output (P*).

The social net benefits under this form of monopoly equal include areas 1+2. Gains to trade are are realized, but additional net benefits are possible. Area 4 would also be realized if production were at Q** instead of Q*. Thus, monopolies of this sort are inferior to competitive markets, because fewer social net benefits are realized than could have been ([1+2] < [1+2+4]). The granting of monoly privileges in such cases reduces social welfare and should be avoided.

Note that the result is similar to that of other barriers to entry, although in this case the price is chosen by the firm, rather one that emerges impersonally as prices adjust to equate supply and demand. Consumer surplus falls relative to competitive markets and prices increase. The reduction in consumer surplus (and profits by excluded firms) would, of course, generate opposition to monopoly polices.[9]

In cases in which competitive markets for the services are possible, both utilitarians contractarians would oppose the creation of monopolies and other laws that enable such markets to emerge. Social net benefits fall and the distribution of net benefits tends to become less equal. Both effects tend to reduce aggregate utility and support for monopolistic markets. Some types of markets are better than others from both perspectives.

The second type of monopolist is the one that is able to sell its output at different prices to different individuals. A discriminatory monopolist is able to sell more of its product, which tends to increase social surplus. However, the distribution of net benefits shifts even more in favor of the owner(s) of the monopoly. In the limiting case, of a perfectly discriminating monopolist, every unit produced is sold at the highest price that consumers are willing to pay. In that case, the monopolist's marginal revenue curve is the same as its demand curve, and output Q** is produced. Social surplus is maximized as in competitive markets, but the division of net benefits is

[9] One of the earliest mass protests in English history was associated with Queen Elizabeth's monopoly policies. Her monopoly grants included markets for iron, salt, starch, lead, paper, vinegar, etc., (Hume, 1688, Brewer 1878). In response to those protests and Parliamentary action, Elizabeth promised to eliminate many of her new monopolies in 1601. As monopoly grants (patents) continued under her successor James I, parliament adopted a stature to regulate future patents and monopolies in 1624, which however, did not fully end the sale of monopoly privileges or their use to reward royal supporters. Town-level monopolies were also commonplace, about which Adam Smith was still complaining in 1776.

completely one sided with monopoly profit accounting for all of the social net benefits.

From the perspective of welfare economics, social net benefits are maximized, so no further improvements are possible. However, from the perspective of contractarian analysis the distribution of net benefits has become less favorable. From behind a veil of ignorance or uncertainty is that they will be a consumer and realize no benefits from this market. Unless, all persons are risk neutral, they are very like to prefer competitive markets to both forms of monopoly. Indeed, they will prefer the first type of monopoly to the perfectly discriminating monopoly. (The same conclusions would follow from a strictly utilitarian analysis, as opposed to the applied form used in welfare economics, because of diminishing marginal utility.)

All three normative analyses thus conclude that some forms of markets are better than others, and that policies that bring about inferior types of markets should be avoided. Commerce produces mutual gains from trade in most monopoly markets, but fewer such gains are realized than are possible in less restricted markets and gains to trade are more narrowly distributed.[10]

D. Pre-Neoclassical Assessments of Policies that Promote Monopoly

It bears noting that the neoclassical case against monopoly was not the first developed. Pre-neoclassical critiques emphasized the same efficiency and distributional concerns as noted above, but without the geometric or mathematical models to ground their arguments. For example, both La Court (1662) and Smith (1776) make critical comments regarding monopolistic markets that were for the most part created through public policies.

> [**N]o members of those guilds**, under what pretext soever, **can be countenanced or indulged in their monopoly, or charter**, but by the excluding of all other inhabitants, and consequently to the hindrance of their country's prosperity. **For how much soever those members sell their pains or commodities dearer than if that trade or occupation was open or free, all the other better inhabitants that gain their subsistance immediately, or by consequence by a foreign consumption, must bear that loss**. [Pieter de la Court (1662) *The True Interest and Political Maxims of the Republic of Holland.* Liberty Fund (p.78).]

> **People of the same trade seldom meet together, even for merriment and diversion, but the conversation ends in a conspiracy against the public, or in some contrivance to raise prices**. It is impossible indeed to prevent such meetings, by any law which either could be executed, or would be consistent with liberty and justice. But though the law cannot hinder people of the same trade from sometimes assembling together, **it ought to do nothing to facilitate such assemblies;** much less to render them necessary. [Smith, Adam (1776 / 2010-03-23). *Wealth of Nations: Full and Fine Text of 1776 Edition* (p. 90).]

Proponents of commerce have long opposed to government policies that create or encourage monopolies, with the possible exception of temporary ones (patents) for new inventions. They did so based in part on self interest and in part from the perspective of pre-utilitarian ethics.

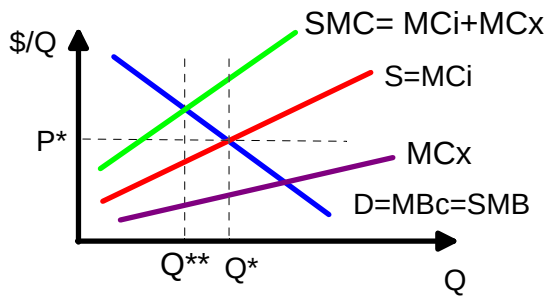
Support for commerce did not emerge with utilitarianism and is rarely indifference to the kind of manner in which commerce takes place. Open competitive markets, ones consistent with principles of equal liberty, have long been preferred to closed ones dominated by privileged towns, firms, or families.

[10] An exception to this rule is the case of a firm that brings a new product or service to market, where the increase in opportunities tends to create new gains to trade. Such firms usually are only temporary monopolists. However, there are a few cases in which such firms continue their monopoly position--not through cartel agreements or special legislation--but by providing exceptional services to their customers and generally assume that such spillover costs and benefits are ignored by firms and consumers in the markets producing them.

IV. Market Failure: Externality Problems

Normative theories also have policy implications with respect to product and service markets in which the production or consumption of goods and services generate costs or benefits for others outside the market of interest. Economists refer to these external costs and benefits as "externalities." Although many externalities were matters covered by civil law, neoclassical economics provided a new clear statement of the nature of the problem and together with welfare economics suggested new solutions. The proposed solutions arguably would make commercial societies more attractive than they would have been under civil law alone.[11]

Figure 10.5 Market Equilibrium and Externalities



The maximize social net benefits norm requires all costs and benefits to be taken into account. If their are costs or benefits not accounted for by market participants, it is likely that market equilibria will also fail to account for those costs. If so, markets will fail to maximize social net benefits for such products and services.

Figure 10.5 illustrates the standard neoclassical representation of an external cost such as air, water, or noise pollution generated by the production (or consumption) of a good or service. The key assumption is that part of the cost of production (or consumption) falls on third parties who do not participate in the transactions that generate them. If firms and consumers are narrowly self interested, they will simply ignore the existence of those "spillover" costs. This assumption implies that all the persons of interest are what has been termed pragmatists in part II. It suggests that the persons engaged in the transaction are not inhibited by their own ethical norms from imposing costs on their neighbors.[12]

Given that assumption, the decision-relevant costs and benefits are captured by the demand and supply curves as above, but the external costs (or benefits) are not. The demand curve can again be used to approximate the social marginal benefits associated with the good or service. However, the supply curve can no longer be used as an approximation of the social marginal cost of producing the good or service of interest. The external costs (MCx) have to be added to industry marginal costs (MCi) to determine social marginal costs (SMC). In a competitive markets, the equilibrium is at the price that equates the quantity supplied with the quantity demanded, Q*. The ideal level from the perspective of welfare economics is Q**, which in this case is somewhat lower than that generated by markets. From that same perspective, there is a problem. Markets in such cases fails to maximize social net benefits.

Insofar as social net benefits correspond to aggregate utility, utilitarians would favor government interventions to improve the result generated by the market. For example, they would support law suits in civil law (torts) to recover damages from the external costs. Or, in cases in which civil law remedies are insufficient to produce optimal results, they might favor excise taxes on the externality generating product. A tax equal to the marginal spillover damages at Q**, will induce market participants to change their behavior in a manner that produces exactly the social net benefit

[11] Ronald Coase (1960) argues that the concept of an externality is itself ambiguous, and moreover that exchange under civil law can solve a variety of externality problems. With respect to the definition of externality, ambiguity can be reduced by taking a temporal account of spillovers. With respect to trade as a solution, as noted by Coase, transactions costs often rule such market solution out. In such cases, he suggests that the civil law should provide a system of rights that produces outcomes that are relatively close to optimal.

[12] There is some recent survey evidence that suggests this is true on average, but not for all people (Fitzgerald xxxx).

maximizing output. (This tax has come to be know as a Pigovian tax.) Pigovian taxes induce firms and consumers to "internalize" the external costs generated by the production of their product.[13]

The benefits of commerce are taken for granted and included in the analysis. The recommendation is not to eliminate the externality, but to take it into account. The policy advice attempts to improve the effectiveness of markets at issue, rather than eliminate them. Such policies would, it is argued, increase the net benefits generated by markets.

A. An Extreme Case of Positive Externalities: Pure Public Goods

The case against negative externalities is clear from a variety of ethical perspectives. In a civil society, individuals should not take actions that harm others. Indeed, other ethical systems often conclude that such actions should not be taken at all, regardless of the individual benefits realized by such actions. Welfare economics suggests that such market activities are OK, as long as the marginal costs imposed on others are less than the marginal net benefits realized by those undertaking the market activities.

In contrast, behavior that generates positive spillovers would under many ethical theories be praiseworthy and often are instances of virtue. Acts of bravery and liberality, for example, often generate benefits for others at the same time that they contribute to character formation. The welfare economics analysis of such activities suggests that they tend to be under done, although this conclusion ignores the effects of both ethical dispositions and praise that tend to increase those activities.

Welfare economics implies that somewhat more of these beneficial activities should be engaged in, a conclusion that is consistent with theories of virtue that argue that more virtue is always better than less virtue. However, in contrast with those virtue-based theories, welfare economics implies that there is an optimal level of external-benefit generating activities, namely that where the social marginal benefits from the activity equals its social marginal cost.

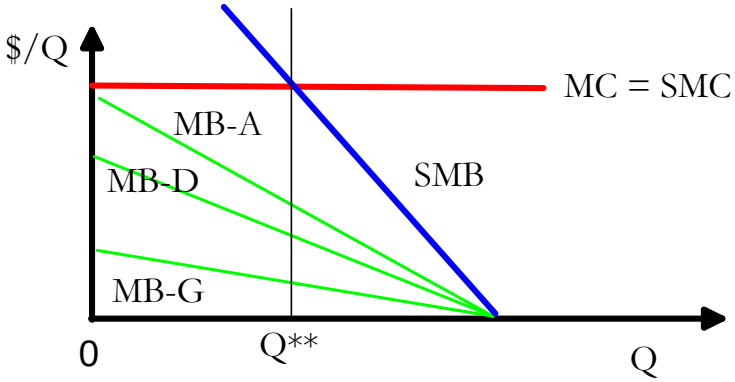
The production of pure public goods are an extreme case of a positive externality generating goods or services. When a unit of a pure public good is provided, everyone in the community of interest receives benefits (Samuelson, 1954). In some cases, each person's marginal benefits from the good may be below their marginal cost over the entire range of interest. In such cases, markets not only fail to maximize social net benefits, they fail to produce the product or service. That textbook case is illustrated in figure 10.6.

Three persons are assumed to benefit from the good or service in question. Yet, each of their marginal benefit curves lies below their marginal cost curve (the full marginal cost or price of provision) over the entire range of interest. Given this, each person chooses not to purchase or provide the goods or services. If one adds all the three private marginal benefit curves to create a social marginal benefit curve, that sum lies above the good's or service's marginal cost curve. If that marginal cost curve is used to approximate the social marginal cost curve, the ideal social-net benefit level can be characterized, which is labeled Q**. As in the previous externality diagrams, private choices lead to "sub-optimal" outcomes in terms of the social net

[13] Spillover benefits, in contrast to spillover costs, have no routine remedy in civil law. There is no invasion of rights, no trespass, no damages. Welfare economics, nonetheless, argues that the cases of external costs and external benefits are fundamentally similar. In such cases, the demand curve would no longer measure all relevant marginal benefits. External marginal benefits would be added to those of consumers to generate the social marginal benefit curve, which will lie above the demand curve. As a solution, Pigou would recommend a subsidy equal to the external marginal benefit at Q**.

benefit norm. Too little--indeed zero--of this good or service is provided in equilibrium.

Figure 10.6: The Public Goods Problem



Welfare economics implies that Q** units of the service should be provided, an in this case, it is often concluded that the service be provide directly by government (or some other tax-financed organization), rather than left to the market. The required Pigovian subsidies would be too large or too difficult to implement.

Public goods, thus, provides another instance in which welfare economics implies that market outcomes can be improved upon by government intervention. In this case the provision of services, rather than regulation or incentive changing taxes or subsidies, is often recommended.

For contractarians, the existence of such problems is one of the justifications for collective or governmental action (Buchanan, xxxx). For utilitarians, it is simply another problem--another market failure--that governments should address.

Government production of public goods enhances life in the community of interest without direct effects on commerce. If government purchases the services, rather than produces them, the solution tends to extend commerce by incentivising firms to undertake the production of goods and services that would not exist in ordinary markets. In the latter case, solutions to public goods problems may be said to extend the commercial society.[14]

Funding the public service, however, requires taxes of one kind or another, most of which have the effect of reducing the extent of commerce in other areas. Through these fiscal effects, public goods problems create a new trade off between the quality of life in society and the extent of commerce.

B. Democratic Failures

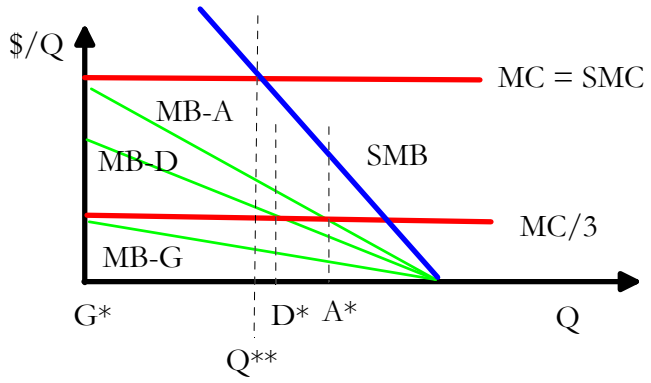
Of course, it is one thing for moral theorists or welfare economists to recommend to government officials or voters that monopoly or externality problems be solved through public policy and another whether such policies adopted. The adoption of policy in any government is an act of choice by one or more government officials. Even a well-functioning democracy is otherwise unlikely to adopt or implement the policies recommended by welfare economics, unless the relevant policy makers have also internalized the social-net-benefit maximization norm.

To see this, consider the case in which voters will directly choose the level of a pure public good to be provided under a preexisting tax system. To simplify the diagram, assume that that tax system simply divides up the costs among tax payers equally, as is often done in private clubs and cooperatives. Given that tax system, each voter favors the service level that maximizes his or her net benefits. By assumption they each confront the same marginal cost (MC/3), but differences in their marginal benefits imply that their ideal service levels differ. Differences in marginal benefits may reflect differences in income, tastes, ethical dispositions, health, and time available for leisure.

[14] In many cases, the firms producing public goods are monopolists or near monopolists, which creates bargaining problems between the government and their producers. Bargaining in that environment can be quite complex and not likely to maximize social net benefit from the service provided. See, for example, Laffont and Tirole (1993).

Figure 10.7 illustrates the ideal service levels of three pragmatic voters, Anthony, Duncan, and Gordon who share the costs of a public good or service. These three ideal points can be used to illustrate the process of voting among net-benefit maximizing persons. Voters with linear marginal benefit and marginal cost curves act as spatial voters, that is to say they will vote for the alternative that is closest to their own ideal point. As drawn, Anthony's ideal point is the highest of the three (A*), followed by Duncan's (D*). Gordon prefers 0 over the other possibilities under the assumed equal cost-sharing arrangement, because of his lower marginal benefits.

Figure 10.7 Selection of a Public Service Level via Majority Rule



Note that in pair-wise elections, D* can defeat any other alternative. Duncan is the median voter, the voter whose preferred policy is the median of the distribution of voter ideal points. Anthony and Duncan will vote for D* over any service level below D* and Gordon and Duncan will voter for D* over any service level greater than D*. If D* is on the ballot, it will be the outcome. The same logic can be extended to any odd number of such voters in a one-dimensional choice setting.[15]

The welfare economics used above to evaluate the private supply of public goods can also be used to evaluate median voter outcomes. The social marginal benefit (SMB) curve is the vertical sum of the individual marginal benefit curves. If the marginal production cost curve (MC) includes all relevant costs, then it can be used as the social marginal cost curve. Together these assumptions imply that social net benefits are maximized at Q**, which in this case is a bit less than the output favored by the median voter (D*).[16]

Note that majority rule may be said to have "failed" in the same sense that the private provision of public services did. Majority rule fails to maximize aggregate utility or net benefits. Nonetheless, in this case, the result is clearly an improvement over the former in terms of social net benefits. More social net benefits are realized at the median voter's choice (D*) than at the private choice (0).

Other ethical theories can also be used to judge the outcomes of majoritarian decision making. For example, from the point of view of private ethics, some might argue that the effect of this policy on Gordon is unethical. He was quite content with the original result and the new one makes him worse off, because of his tax obligations. Contractarians support unanimous agreement, when possible, and would insist on compensating Gordon for his losses or adopting another tax system.[17] For example, contractarians often favor Lindahl tax system, which divides costs according to individual

[15] For more on the median voter and median voter theorems, see Congleton (2003) or Mueller (2003).

[16] If public production generated negative externalities, that the social marginal cost curve would be higher than that curve, and so the optimal level of the public service would be lower. The early tests of nuclear weapons, for example, left large tracts of radioactive land and generated significant fallout. Military aggression imposes costs on persons living in the area under dispute. Highway, metro, and office construction generates noise and congestion. We ignore such nontrivial externalities here to focus on decisions comparable to those analyzed in chapter 10.

[17] James Buchanan (1987) often made this point, insisting on what he called "Wicksellian unanimity" as the proper way to choose public policy. For a collection of several papers on how coercion aspects of taxation affect normative conclusions about taxation, see Martinez-Vazquez and Winer (2014).

marginal benefits at Q**, would generate unanimous support for the social net-benefit maximizing service level.[18]

In general, other normative systems may reach quite different conclusions about the level of a pure public good, because they focus on other consequences or dimensions of life than social net benefits or aggregate utility.

Differences among utilitarians and contractarians would also be introduced by differences in the assumptions about human nature or the distribution of ethical dispositions in the community of interest. A magnanimous person might for example, provide Q** units of a public good or service to his or her community, because of a sense of duty or virtue, or because of the expected approbation associated with that behavior. If such persons exist in a community, public goods such as concert halls and libraries might well be provided without collective action. The extent of public goods and externality problems vary with the distribution of ethical dispositions in the community of interest.

In a community of pure utilitarian, for example, each person would attempt to fully account for such effects in all of their actions. Externality problems would only exist by mistake. Many other civic and reciprocity norms also require externalities to be taken into account. Thus it is likely that at least some firms and consumers would undertake the required internalization of externalities without the intervention of law. Such ethical consumers would be willing to pay a premium for products produced in a manner that took externalities into account. If a sufficient number of such consumers exist, market forces would induce internalization, rather than undermine private efforts to do so.

The neoclassical analysis is a worst-case analysis. Externality problems in many communities would be smaller than neoclassical models imply and public policies better.

V. Improving the Market-Based Distribution of Income

Another area of concern raised by twentieth century welfare economists and utilitarians is the distribution of income generated by commercial societies. In their positive research, economists generally take the distribution of wealth and income to be facts generated by "nature." They are (largely) determined by market forces generated by inheritance, tastes, competition, and technology. Within a perfectly competitive market wages reflect the marginal revenue products of the persons employed, which is to say the value that they add to the social dividend.[19] The resultant inequality of wages reflects differences in prices of the outputs produced and in the marginal products of individuals and teams producing those goods and services. Some persons are stronger, smarter, better educated, and entrepreneurial than others. Some talents and ethical dispositions are more scare than others, and so some careers are better paid. Some persons choose to specialize in areas in which unanticipated innovations increase or diminish their marginal value products.

Differences in the distribution of wealth is largely a consequence of similar market forces, given differences in inheritance and saving rates. Additional inequalities are produced by good fortune in one's investments, monopoly power, and government favor. Markets for assets and labor are just like any other insofar as prices are determined by supply and demand.

Philosophical concerns about a society's distribution of wealth predate the emergence of utilitarianism and the commercial society. Wealth in prior times was largely based on land holdings and largely inherited. Distributions of wealth within towns and cities were also largely based on inherited status and trade privileges, although more was determined by market forces, because commerce was more central to life. Land holdings outside cities were generally very difficult to transfer or sell to persons outside one's

[18] Under a Lindahl tax system each person pays a tax (marginal cost share) equal to his or her marginal benefits at Q**. With these marginal cost curves, there would be complete agreement about the optimal service level, namely Q**.

[19] Marginal value product is the price of the output produced multiplied by marginal product. The marginal product of an input is the change in output generated by a one unit change in that input, e.g. the output generated by an additional hour of labor.

family. For example, rights to a rural strip farm was routinely passed on to one's children, but rarely could be sold.[20]

As the family-based land and legal entitlements of the middle ages gave way to market-based ones, theories of distributive justice emerged to assess the merits of the pattern of wealth and income associated with more or less open commercial systems. These normative theories were not simply efforts to rationalize the existing distribution of income, as neoclassical theories may be said to, but to assess their merits relative to a hypothetical ideal. That ideal, naturally, varied according to the normative theory applied.

We begin with an updated version of the theory of distributive justice that dominated the nineteenth century and much of the twentieth century. Namely, that one owned oneself--including all of one's capacities--and one's inherited wealth. And, one should be free to use those resources as he or she sees fit, as long as the rights of others are not violated, where rights for the most part meant the rights defined by existing civil and constitutional law. The resulting distribution of income and wealth was considered both lawful and just.

A. Distributive Justice as a Characteristic of Process rather than Outcomes

The following quote from Nozick (1974) outlines the logic behind this "classical liberal" approach to judging market and other outcomes.

> 1. A person who acquires a holding in accordance with the principle of justice in acquisition is entitled to that holding. 2. A person who acquires a holding in accordance with the principle of justice in transfer, from someone else entitled to the holding, is entitled to the holding. 3. No one is entitled to a holding except by (repeated) applications of 1 and 2.
>
> The **complete principle of distributive justice would say simply that a distribution is just if everyone is entitled to the holdings they possess under the distribution**. A distribution is just if it arises from another just distribution by legitimate means. **The legitimate means of moving from one distribution to another are specified by the principle of justice in transfer**. (1974, *Anarchy, State, and Utopia*, p. 151).

What matters from this perspective is not the global properties of a commercial or social system, but how it emerged and where it started from. Is property initially in the hands of its rightful owner. Are the individual acts that generate the system legal and ethical? If so then the distribution of wealth and income associated with that system may be defended on moral grounds.

Insofar as inherited wealth is legitimately acquired (as a gift from previous generations) and one's subsequent income is acquired through voluntary exchanges with others (who have clear title to what money or other things they trade for the labor and other resources provided), the resulting distribution of income and wealth is both legitimate and moral. In other words, if you have come by your wealth and income honestly and without coercion, your wealth and income are legally yours and rightfully should be.

B. Utilitarian Challenges to the Income Distribution Generated by Commerce

Utilitarians generally accepted this theory until the idea of diminishing marginal utility was introduced and became widely accepted among both utilitarians and economists. At that point, it became clear that redistributing wealth from the rich to the poor could potentially increase aggregate utility. If rich and poor persons had fundamentally similar utility functions, then taking a dollar from a rich person and giving it to a poor person would

[20] This was the norm in Europe for many centuries, until the enclosure movement of the eighteenth and nineteenth centuries. After "enclosures" were adopted--which often required acts of parliament--marketable free hold titles to land became more commonplace.

generate a smaller reduction in utility for the rich person than it generates for the poor person.

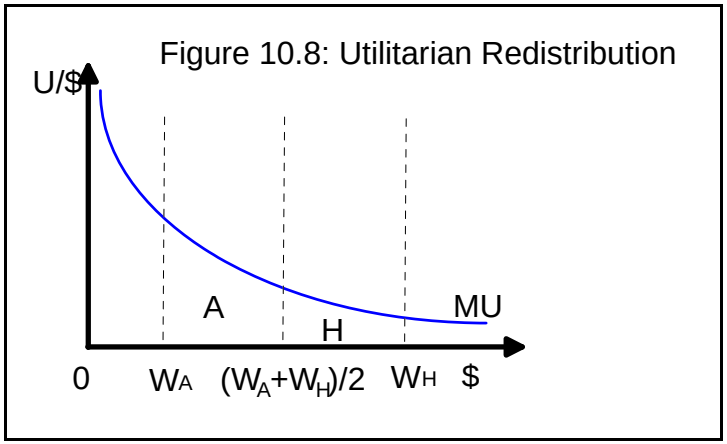


Figure 10.8 illustrates this idea. Suppose that Alfred and Henry have identical utility functions and therefore marginal utility curves. Assume that their utility functions exhibit diminishing marginal utility, which means the successive units of a good generate less and less additional utility. Logically, each will use his money to purchase the most valuable (utility generating) goods first, the second most valuable good second, and so forth. Now, assume that Alfred and Henry have different levels of wealth, W_A and W_H. Alfred's total utility is the area under the marginal utility curve from 0 to W_A and Henry's is the area under the marginal utility curve from 0 to W_H.

Notice that if each person had the same income ($[W_A+W_H]/2$), that aggregate utility increases. Henry loses area H in utility but Alfred gains area A in utility. A > H, and thus the sum of their utilities increases. Aggregate utility rises by H-A.

Once understood, this idea altered the utilitarian perspective on commerce. Although a good system, commerce was no longer necessarily the best system that could be imagined. Debates among utilitarians thenceforth shifted to the feasibility of redistribution schemes like that involved in Figure 10.8.

Although it may be physically possible to equalize income, would the policies needed actually increase aggregate utility. How would redistribution affect the amounts available for redistribution? Surely if all persons were paid the same amount regardless of how hard they worked, very few would work hard. Moreover, if a political system were given the authority to undertake such major shifts in wealth and income, what policies would in fact be adopted? Would a government with such authority tend to increase or diminish aggregate utility?

Economic tradeoffs between equality of income and total income were the main focus of the debates among utilitarians and welfare economists. The political problems were little discussed by welfare economists. They are discussed in chapter 11. The new debates arguably tended to undermine ethical support for commercial systems. Although they had relatively little effect on public policy until the second half of the twentieth century, they tended to challenge one of the most obvious consequences of a commercial society.

C. Distributional Tradeoffs: Redistribution, Production, and Welfare

The egalitarian solution to the Edgeworth allocation problem is consistent with utilitarian analysis if the resources available are not affected by the manner in which they are distributed (and preferences are identical). If, however, the size of the Edgeworth box is determined by the manner in which resources are allocated, these incentive effects have to be taken into account. For example, if some or all of the goods and services of interest are produced with labor, the effects of alternative distributional mechanisms on the labor supply have to be taken into account.

Figure 10.9 illustrates some of the tradeoffs between income equality and total income at the level of a firm. Suppose preferences are identical for leisure (L) and a single produced output (C) that can be produced either alone or by teams. Assume that a subset of the necessary jobs on teams are unpleasant and that every job on the team is less pleasant but more productive than that of solo production. In the absence of an internalized work ethic or biological necessity, no one would work for the joy of working.

Several more or less equal payment systems can be analyzed in this relatively simple production environment. The output and salary effects of

four hypothetical rules for distributing a cooperative's or firm's output among its members are illustrated in figure 10.9.

Two egalitarian rules and two non-egalitarian rules are represented. Under system 1, *equal shares*, output is always equally shared among all *N* persons in the community. This payment system satisfies utilitarian norms in settings like that illustrated in figures 10.8 and 10.1. In this setting, however, the goods and services to be distributed must first be produced. Each potential laborer chooses whether to work or not. Under the equal sharing rule, they will all free ride (shirk rather than work). They gain only 1/N of the output produced by their own efforts, but realize all the benefits of their leisure. They will work only if the value of 1/N of their output is greater than the value of the leisure that has to be sacrificed to produce it.[21] This is unlikely in the absence of biological necessity. Thus, everyone free rides, and the community produces no output (above subsistence). In effect, such a sharing rule imposes a (N-1)/N tax rate on each person's wage.

Under system 2, *equal wages*, everyone in the community is paid the same hourly wage for their efforts, W, regardless of the type of firm or task. Hours of work are rewarded, but not the difficulty of it. Thus unpleasant tasks would not be done, and team production would be less productive than it could have been. Indeed, under the assumptions, everyone would work for themselves, because this is a more pleasant mode of production than team production. This wage scheme reduces the incentive to free ride, because shirkers will be paid zero for their lack of efforts (W times 0 hours of work). It does not, however, induce the most productive forms of production, because they are assumed to be unpleasant. If NL* hours are worked in total and NQ* units of output are produced, the wage rate can be as high as Q*/L*, the average rate at which output is produced per hour by a typical one-person team in the community.

Under system 3, *equal wages within firms*, wage rates may differ among firms, but are uniform within each firm. Multi-person firms can now pay team members more than they could make in solo production, which can induce persons to join economic organizations and participate in their unpleasant modes of team production. Because of the assumed productivity differences, wage rates will be higher at the factories than for solo production.

All labor will be attracted to the factories if the productivity advantage of team production is sufficient to more than compensate persons for their less pleasant working conditions. The higher wage rates would not, however, induce persons to invest in difficult skills or engage in the most unpleasant tasks within firms.[22] In such cases, community output increases over that of system 2, and average income increases, which raises utility for the average person in the community. The social dividend is smaller than it could be, although everyone reaches a higher indifference curve.

(4) The fourth system, *competitive wages*, is similar to that observed in commercial systems. Compensating differentials are paid within firms for more or less pleasant work and for productivity differences associated with talent, training, and intensity that individuals brings to their teams. In this case, a better (more productive) allocation of persons to specialties within the factory takes place and output rises again. As total output increases, average wages increase, which increases average utility. This in turn tends to imply an increase in both average and aggregate utility, although the distribution of income and wealth is far from equal under this system.[23]

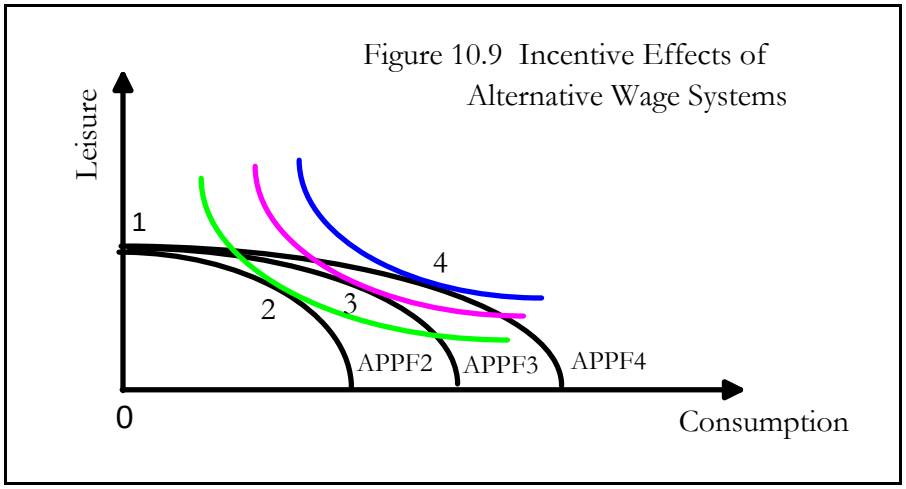
---

[21] Other equilibria are possible, although all involve relatively low levels of output in the absence of counter-veiling private norms such as personal work ethics..

[22] In a more general treatment, specialized suppliers of unpleasant tasks or highly specialized tasks would tend to emerge under this egalitarian wage system to reduce such problems. Inequality among employees of different firms would emerge in a more fully developed egalitarian wage system. Willingness to invest in human capital or tackle the unpleasant tasks would be preconditions of employment at such firms. However, intra firm production of services, as in system 4, are evidently more efficient than subcontracting for such services.

[23] If average utility increases, then aggregate utility necessarily increases, $W = N U^{ave}$. Whether average utility increases over that generated in system 3, depends on distributional issues as well as the productivity effects of productivity oriented wage systems. Aggregate utility tend to rise over system 3, if team member / employees in the middle range of talent and work effort earn approximately the average salary, as in figure 10.9. On the other hand, if

Figure 10.9 illustrates the average person's choices and welfare under these four regimes. An average production possibility frontier (APPF) is created by dividing every point on the community's production possibility frontier by N, the number of residents in the community. Each wage system has associated with it a different (average) production possibility frontier (APPF) reflecting how the distributional rules affect the mode of production and labor-leisure choices. Along any average production possibility frontier, increases in the consumption good require reductions in leisure (increases in the hours worked). The average PPFs allow a middle-class person's indifference curves to be used to illustrate how wage and transfer systems affect aggregate utility.



It is important to note that the egalitarian utilitarian solution to the distributive justice problem (system 1) fails to maximize aggregate utility. Whenever incentive matters, utilitarian analysis take incentives effects into account. In the illustration, the competitive wage system dominates the other three.

The distribution of income is equal among the first two cases, but unequal in the last two. Both systems 3 and 4 tend to dominate the egalitarian divisions of aggregate output under utilitarianism. Under system three, wages vary among employees of specialized firms and within system 4 among persons within each firm, because of differentials paid to compensate for special skills and the unpleasantness of the various jobs associated with specialized forms of team production. The commercial system (4) dominates the regulated wage system (3) if productivity gains within firms yield sufficient additional material comfort (higher average wages) to offset the diminishing marginal utility effects of greater wage dispersion.

VI. Contractarian Theories of Distributive Justice

The contractarian methodology for assessing the distributions of income associated with alternative social systems differs from the utilitarian one. Contractarians imagine the decisions about alternative systems to be reached by individuals, rather than a utilitarian social planner. They believe that all (or essentially all) individuals would reach the same conclusions about the systems considered, or at least concerning the main features of such systems from behind a veil of uncertainty or ignorance. Contractarians also tend to focus on system wide choices and reforms of such systems rather than day to day policy choices or personal conduct.

There are also points of agreement among utilitarians and contractarians. Both are consequentialist theories and thus economic analysis also plays a central role in contractarian assessments of the income distributions associated with commerce. It is anticipated consequences that allow principles of justice and institutions to be ranked by individuals. This is true under the Rawlsian approach where common principles of justice are developed, and also under the Buchanan approach where institutions are selected. There are, however, points of disagreement among contractarians.

---

sub-contracting is nearly as efficient as internal management and control, system 3 may generate higher aggregate utility than system 4. (It bears noting that a good deal of the lower income inequality in Scandinavia is generated by their national rules for wage rates, which are somewhat in the spirit of system 3.) The diagram assumes that wage dispersion within competitive wage system is not too different from that generated in a system 3, and team production increases substantially in efficiency-- by reducing transactions and information costs.

Figure 10.9 can be used to illustrate how contractarians analyze the distribution of income associated with commerce. From a Rawlsian perspective, individuals would reach agreement about the principles of justice that would be used to evaluate alternative institutions. They would subsequently attempt to determine which of the feasible systems is most consistent with those principles. From the difference principle alone, Rawls would anticipate that system 1 would be rejected. Every one is better off in systems 2, 3, and 4 than they are in system 1. With respect to the maximal equal liberty principle, system's 3 and 4 arguably provide greater liberty than system 2, insofar as more variety in career paths are likely under those systems than system 2. Whether, there is a tension between the maximal equal liberty principle and the difference principle, depends on whether the lowest wages in system 3 or 4 lie below that of system 2. This is possible, but less likely than might be supposed given the productivity advantages generated by specialization. With respect to comparisons between systems 3 and 4, again productivity differences and their effects on the lowest wage rates are important. Liberties are nearly equal under the two systems, although somewhat broader under system 4 than system 3, because more forms of voluntary association are allowed under 4 than 3 (e.g. firms within which wages vary).

From the Buchanan perspective, the starting point matters. If community is initially in system 1, then Buchanan would anticipate unanimous agreements to adopt rules compatible with systems 2, 3, or 4. Given that average material welfare increases as one moves from 1 to 2 to 3 to 4, he would argue that system 4 would be the most likely choice, unless the people of interest are very risk averse. The more risk averse persons in the community are, the more willing they are to accept a decrease in average income for a decrease in wage dispersion. Thus, it is possible that a community would unanimously agree to 3 in some cases.

On the other hand, if the starting point is system 4, shifts to system 3 would be unlikely, because many persons would anticipate being made off by such a system and vote against it. Without (near) unanimity for a major reform of their existing system, the status quo dominates. A similar conclusion might be reached if the status quo were system 3 instead.

That the commercial societies that emerged in the 19th and 20th centuries resembled system 4, a lack of consensus for revisions, would imply that the commercial system is ideal from Buchanan's contractarian perspective.

All the above analysis of ideal distributions of income are undertaken from an "other things being equal" perspective. In other words, it assumes that shifts in wage systems affect on the extent of leisure and material comfort, not the distribution of ethical dispositions, nor the kind of political system in place. A complete analysis would account for effects of market systems on other objectives such as ethical development and on the political systems through which market relevant policies are to be adopted. Such interdependencies are taken up in chapter 11.

VII. Welfare Economics as the Definition and Correction of Market Failures

Social, as opposed to personal ethics, attempt to rank actions taken by groups of persons, rather than individuals. For utilitarians, the same principles can be used to assess individual, group, and system level performance. Has aggregate utility for the community of interest increased or not? If it has, then the action, policy, or system is a good one. Other normative theories are less encompassing, but may also have implications for both personal and group actions. Contractarian reasoning, for example, can be used to assess the merits of alternative institutions and policies. It may also be used to analyze the subset of personal ethics that address behavior within groups, what has been referred to as civil ethics above (Buchanan xxxx). Personal ethics may also be used to assess the worthiness of policy makers and policies. Insofar as virtuous men and women tend to be leaders under one political system more often than others, some systems may be regarded as better (morally superior) to others.

This chapter has shown how utilitarian and contractarian analysis can be used to assess commercial activities and systems. It also demonstrated how the models employed produce recommendations for reform, if improvements appear to be possible. Some of the analyses provided additional--argueably deeper--support for commerce than that provided by earlier assessments of the merits of markets. Other analysis raised new questions

about the ethical appeal of some forms of markets and market systems as a whole.

Welfare economics implied that the market systems that emerged under 19th century civil law could, in principle, be improved by reducing monopoly power and addressing externality and public goods problems. Other larger issues were raised about the consequent distribution of income were raised and used to justify interventions in labor markets and new redistributive and social insurance programs.[24]

It bears keeping mind that although the analysis often was undertaken with sophisticated mathematical models and included insights from neoclassical economics, all the "market failings" identified were consequences of the ethical theory applied. Welfare economics has explicitly utilitarian foundations. Contractarian analysis departs from utilitarian ideas, but uses unanimity and voluntariness as norms for ranking policies and institutions. Other normative systems, of course, could also have been applied and may have generated different conclusions.

Many of the problems identified by twentieth century welfare economists had been noted in earlier centuries. There had been public demonstration against monopoly since at least 1600. And, although the term externality originated in the twentieth century, no one in earlier times defended a gun owner's "right" to use his neighbor's house for target practice. Ethical theories and dispositions had long played central roles in the identification of "problems," in proposals for solutions, and in critiques of public policies. The new methodology of applied utilitarian analysis simply made the arguments sharper and often provided new quantitative measures of the extent of the problems identified.

It also bears keeping in mind that the new analyses of market failures and public policy remedies were not themselves without failings. For example, the informational requirements of both utilitarian and contractarian analysis are enormous. Individual preferences have to be completely understood. The consequences of policies, reforms, and institutions have to be accurately assessed. Both require more complete and accurate models of economic development and better data than presently exist. General trends in macroeconomic aggregates are fairly easy to measure, but models and estimates of the activities that generate those aggregates are far less precise, and it is these which directly affect individual welfare, rather than macroeconomic aggregates

There are deeper issues as well. It is one thing to suggest that utility maximizing man and competitive markets are useful analytical constructs that shed light on how well-developed markets operate, and another to say that they represent exactly how people reach decisions and how markets operate. Moreover, it a variety of interdependencies clearly exist among ethical dispositions, the extent of commerce, public policies, and social systems, many of which are neglected in textbook treatments of the need for and effects of public policies. All these imply that the mathematics and geometry of contemporary normative analysis tends to exaggerate the precision of the conclusions reached.

With these interdependencies in mind, the next chapter analyzes the role of ethics in the political and legal institutions that frame commerce and explores some institutional implications of that analysis.

References

Berggren, N. and Jordahl, H. (2006) "Free to Trust: Economic Freedom and Social Capital," *Kyklos*: 141-69.

Buchanan, J. M. (1987) "The Constitution of Economic Policy," *American Economic Review* 77: 243-50.

Coase, R. H. (1960). "The Problem of Social Cost," *Journal of Law and Economics* 3: 1-44.

[24] The most common result of studies that examine the effects of "economic freedom" on indicators of human welfare is that economic freedom increases human welfare. See, for example, Gwartney, Lawson, and Holcombe (1999), de Haan and Sturm (2000), or Berggren and Jordahl (2006).

Congleton, R. D. (2003). "The Median Voter Model" in *The Encyclopedia of Public Choice*, C. K. Rowley and F. Schneider (Eds) Kluwer Academic Press.

de Haan, J. and Sturm, J. E. (2000) "On the Relationship Between Economic Freedom and Economic Growth," *European Journal of Political Economy* 16: 2015-41.

Fitzgerald, M. P.

Gwartney, J. D., Lawson, R. A., and Holcombe, R. G. (1999) "Economic Freedom and the Environment for Economic Growth," *Journal of Institutional and Theoretical Economics* 155: 643-663.

La Court, P. (1662/1746) *The True Interest and Political Maxims of the Republic of Holland.* (Trans.) J. Campbell. London: J. Nourse.

Laffont, J. J. and Torole, J. (1993) *A Theory of Incentives in Procurement and Regulation*. Cambridge Mass.: MIT Press.

Martinez-Vazquez, J. and Winer, S. (Eds) (2014). *Coercion and Social Welfare in Public Finance: Economic and Political Perspectives.* Cambridge: Cambridge University Press.

Mueller, D. C. (2003) *Public Choice III*. Cambridge: Cambridge University Press.

Nozick, R. (1974/2013). *Anarchy, State, and Utopia*. New York: Basic Books. (Kindle Version)

Pigou, A. C. (1920 / 2013). *The Economics of Welfare*. London: Macmillan and Co. (Kindle Version).

Samuelson, P. A. (1954). "The Pure Theory of Public Expenditure," *Review of Economics and Statistics* 36: 387-89.

Smith, A. (1776 / 2002). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Public Domain Books. Kindle Edition.

Varian, H. R. (1992). *Microeconomic Analysis*. New York: Norton.

Appendix: The Net Benefit Maximizing Model of Man and Markets

Rational choice based models for the most part use the utility maximizing model which has utilitarian origins, developed in chapters 5 and 9. Pigou's (1920) classic work on welfare economics suggests using dollar-based measures of welfare. These, as it turns out, can be used to generate an alternative rational-choice based foundation for much (although not all) of neoclassical economics and welfare economics. It is one that is generally easier to apply and provides sharper implications than utility-based analysis. It is that perspective which is summarized in this appendix and used to motivate the area-based welfare economics used in the diagrams of this chapter. To shorten the space required for a quick review, a bit of calculus is employed, and the geometric equivalents are discussed, rather than illustrated, for the most part.

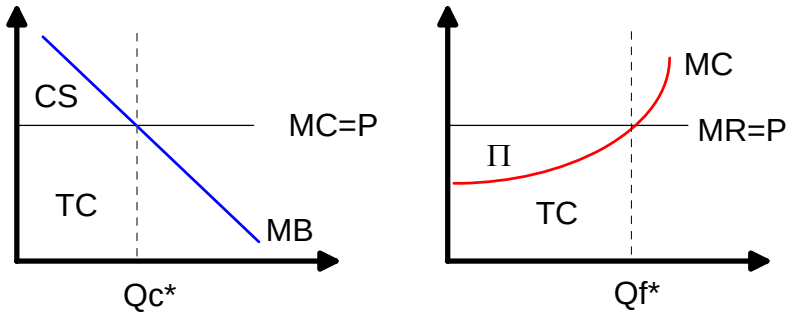
The net benefits associated with purchasing some quantity of good A can be written as: N = b(Q) - c(Q). For consumers, b is the benefit (B) function that describes the highest price in dollars or some other currency that one would be willing to pay to have Q units of good A rather than zero, and c is the cost (C) function that describes how much one actually has to pay to have Q units of the good. For firms, b is the function that determines how the revenues associated with the production and sale of good A rise with the quantity sold and c is the cost function describing the total cost of producing and selling the good of interest.

All these functions include other variables than those written in this short form, as for example function b for consumers includes personal income (Y), the prices of other goods (Po), health (H), and in some cases ethical norms (E), as with B = b(Q, Y, Po, H, E). The cost functions for firms include the prices of inputs (Pi) and production technology (T), and some cases entrepreneurial skill (S) and ethics, (E), and firm location (L), as

with C = c(Q, Pi, T, S, E, L). In order to cut down on notation, all the variables except Q are suppressed in the discussion below. The textbook illustrations focus on what happens in the quantity domain of consumer and firm choice.

Calculus implies that the quantity that maximizes net benefit, Q*, is that where dN/dQ = dB/dQ - dC/dQ = 0. In economic terms, net benefits are maximized at a quantity where marginal benefits (dB/dQ) equals marginal cost (dC/dQ). This is true for both rational consumers and rational firms, although the meaning of the marginal benefit and marginal cost functions is slightly different for the two sides of the market. Firms are often assumed to be interested in a net benefit called profit and in that case, the marginal

Figure 10.10 Consumer and Firm Choices



benefits are represented as the firm's marginal revenue (the amount of extra revenue generated by selling one more unit of the good of interest).

It bears noting that if one knew the marginal benefit and marginal cost functions, one could generate the total cost and total cost functions by taking the integral from 0 to Q of the marginal benefit and marginal cost functions, $b(Q) = {}_0\int^Q dB/dQ\ dQ$ and $c(Q) = {}_0\int^Q dC/dQ\ dQ$. In geometric terms, this means that the total benefit of quantity, Q, can be calculated from a marginal benefit curve by finding the area under the marginal benefit curve between 0 and Q. Similarly, one can find the total cost of quantity Q (actually total variable cost) by finding the area under the marginal cost curve from 0 to Q.

This area principle plays a central role in most of the diagrams in this chapter. Net benefits can be calculated by finding the difference between the areas that represent total benefits and total costs. For readers without calculus this area principle is sufficient to generate all relevant comparisons of consumer, firm, and social welfare.

Next we need to show that demand and supply curves can be used to characterize the marginal benefits of firms and the marginal costs of industry. Demand curves for consumers can be derived by assuming that their cost function is simply C=PQ, where P is the price of the good purchased and Q is the quantity purchased. Notice that in this case consumers purchase amount Q* where their marginal benefits (dB/dQ) equals their marginal cost, which in this case is the selling price of the good. Supply curves for a firm can be derived by assuming that the benefit of production and sale of goods is simply the revenue generated, with R = PQ. Firms will produce amount Q* where their marginal cost (dC/dQ) equals their marginal revenue, which in this case is the selling price of the good.

Figure 10.10 illustrates a typical consumer's and firm's net benefit maximizing choice for the case where they can purchase or sell as much of the product of interest (here A) at the prevailing market price P. Their net benefits are found as areas between the their respective marginal benefit and marginal cost curves, because after subtracting total cost (TC) from total benefits (TB), it this area which is left. The net benefits of consumers is normally called consumer surplus (CS) and that of firms or firm owners is called profit (Π).

Notice that market prices mean different things to consumers and firms. For consumers it represents marginal cost and for firms or firm owners it represents marginal benefits. Note also that a typical consumers benefit from purchasing goods (by areas CS) and a typical firm also profits from selling goods (by area Π).

These diagrams characterize one point on the consumers demand curve and one point on a typical firm's supply curve. When the consumer confronts price P, he or she will purchase quantity Q_c*. When a typical firm confronts the same price, it will produce and sell quantity Q_f*. Geometrically a demand curve can be traced out by varying price, and using the associated marginal cost to find the net benefit maximizing quantity.

For downward sloping MB curves, these will always be found where the marginal benefit curve equals the prevailing price. Thus, it turns out that every point on a persons demand curve is also a point on his or her marginal benefit curve.

Geometrically ,a firm's supply curve is generated in the same way. One varies price and uses the associated marginal revenue curve to find the profit maximizing output. This occurs where marginal cost equal price. So, it turns out that every point on a firms supply curve is also a point on its marginal cost curve.

Thus, when we add up the demand curves of individuals in a particular market (horizontally, in the Q dimension) to form a market demand curve, the market demand curve retains the properties of marginal benefit curves for those consumers. Similarly, when we add up the supply curves of individual firms (horizontally), the market supply curves have the properties of an industry's marginal cost curve. Thus, intuitively, the area rules can now be used to determine net benefits for all firms and all consumers in the market of interest.

To show the mathematics behind this intuition we need to return to calculus. Calculus implies that the ideal quantity for consumers, Qc*, always has the property that dB/dQ - P = 0 (marginal benefit equals price at that quantity).[25] Applying the implicit function theorem to this equation implies that Qc* is a function of price, as with $Q^d = d(P)$ or if all the suppressed variables are brought back into light, $Q^d = d(P, Y, Po, H, E)$. The implicit function differentiation rule can be applied to show that consumer demand curves derived in this way always slope downward (unlike their utility-based counter parts).

Similarly, we find that the firms profit-maximizing output, Qf*, always has the property that P - dC/dQ = 0 (price equals marginal cost at the optimal output). The implicit function theorem implies that Qf* can be written as a function of price, as with $Q^s = s(P)$ or if all the suppressed variables are brought back into light, $Q^s = s(P, Pi, T, S, E, L)$. The implicit function differentiation rule can be used to show that supply curves derived in this way always slope upward ($dQ^s/dP > 0$).

A market demand curve characterizes the overall quantity sold at given prices. This is simply the sum of the individual purchases at those prices. Denote by "i" an individual demand function: $Q_i^d = d(P, Y_i, Po, H, E_i)$. If their are N consumers in the market, the market demand is simply: $Q^D = \Sigma Q_i^d$ for I = 1..N. Similarly, the market supply of a typical firm j is $Q_j^s = s(P, Pi, T_j, S_j, E_j, L_j)$. If there are M firms in or potentially in the market, the market supply is simply the sum of those supply functions, $Q^S = \Sigma Q_j^s$ for J = 1..M.[26]

Markets "clear" when supply equals demand, which occurs at a price, P*, such that $Q^D(P^*) = Q^S(P^*)$. To find the total benefits of all consumers in that case simply requires adding up all the benefits associated with the amounts individuals purchase at that price, $Q_i^* = Q_i^d(P^*)$. If we start with marginal benefit functions, we know that this is the area under each individual's marginal benefit curve from 0 to Q_i^*, $b(Qi^*) = {}_0\int^{Qi^*} dB_i/dQ\, dQ$. The sum of those benefits is simply: $B = \Sigma\, {}_0\int^{Qi^*} dB/dQ\, dQ$ for i=1, .. N, which can also be written as $B = {}_0\int^{Qi^*} \Sigma\, dB_i/dQ\, dQ$ for i=1, .. N.

Since the marginal benefit function goes through the same points as the demand curves in the usual textbook case with diminishing marginal utility, the first expression is in effect the sum of the areas under the individual demand curves and the second the area under the market demand curve out to $Q^D(P^*)$. The total cost of that quantity is $P^*Q^D(P^*)$, which is a rectangle, and the difference between the benefits realized by consumers and that cost is the usual triangular area under the demand curve (as the MB of all consumers) and above the price line (as the MC of all consumers).

---

[25] There are exceptions to this rule at Q=0 and $Q \Rightarrow \infty$, where MB < MC for all Q or MB > MC for all Q, but these are neglected here to focus on the core neoclassical market relationships.

[26] This is the typical derivation of a Marshallian or Ricaradian short run supply curve. It also can be used to characterize a Ricardian long run supply. Marshallian long run supply differs somewhat, but Marshallian markets are not very common in contemporary commercial societies. It is the Ricardian model that is used in the diagrams throughout this book.

The same logic applies for the calculation of industry profit using the market supply curve to represent industry marginal cost. To find the total cost of all firms requires adding up all the costs (variable costs) associated with the amounts firms produce at that price, $Q_j^* = Q_i^s (P^*)$. If we start with marginal benefit functions, we know that this is the area under an individual's marginal benefit curve from 0 to Q_j^*, $c(Q_j^*) = {}_0\int^{Qj^*} dC_j/dQ\ dQ$. The sum of those benefits is: $C = \Sigma\ {}_0\int^{Qj^*} dC_j/dQ\ dQ$ for j=1, .. M, which can also be written as $C = {}_0\int^{Qj^*} \Sigma\ dC_j/dQ\ dQ$ for j=1, .. M.

Since the marginal cost function goes through the same points as the supply curves in the usual textbook case (with increasing marginal cost), the first expression is in effect the sum of the areas under the individual firm supply curves and the second the area under the market supply curve out to $Q^D(P^*)$. The total revenue of that quantity is $P^*Q^S(P^*)$, which is a rectangle, and the difference between the revenues realized by firms and their costs is the usual triangular area above the supply curve (MC industry) and below the price line (MB or MR of firms).

The above derivation of supply and demand, thus, shows why the various areas associated with a standard market diagram can be used to calculate social net benefits in the absence of externalities, as in figure 10.2, which in turn can be used to reach normative conclusions.[27]

Adding consumer surplus and industry profits gives us the social net benefits from production and sales in the case in which there are not externalities.

In the absence of externalities, the demand curve can be used to approximate the social marginal benefits (SMB) of the good or service of interest and the supply curve as the social marginal cost (SMC) of the good or service of interest. In that case, markets in long run equilibrium tend to produce the outputs that maximize social net benefits, which normally occurs at the output where SMB = SMC. This is, of course, the output identified by the intersection of the demand and supply curves in figure 10.2. Market prices, in turn, divide up the social surplus (an alternative measure of the social dividend).

If their are external costs or benefits, these have to be subtracted or added to get social net benefits. These can be simply added to the market supply or demand curve to derive the social marginal cost or social marginal benefit curves. The same logic about marginal costs and marginal benefit curves applies.

[27] This derivation also removes the hob goblin if the need for Hicksian compensated demand curves to calculate consumer surplus, as required under a utility-based derivation of consumer demand, and also eliminates the difference between long run and short run supply in the Marshallian analysis. The only difference above, it the use of short run and long run marginal cost in the supply curve analysis, which for purposes of illustration simply requires shifting from MC_{SR} to MC_{LR}.