# Fundamental Concepts of Microeconomics 6

## 6.1 Objectives and Methods of Microeconomics

Microeconomics is an impressive (and bold) architecture of thought. Its objective is most ambitious: it attempts to describe, explain and evaluate almost everything that is going on in the world as far as human behaviour is concerned.[1] To achieve this goal (or to at least come as close as possible) microeconomics takes a multi dimensional approach providing

- a theory of rational individual decisions,
- a theory of conflict resolution and coordination,
- a theory of the evaluation of resource allocation,
- a theory of public regulation.

Given this set of dimensions, microeconomics is fairly *comprehensive*. It is also quite *general*. Let us see what the term "general" means in this context, and take the first of the four bullet points presented above, the dimension of individual decision making, as the example for our explanation.

The decisions of a baker are certainly different from the decisions of the manager of a basket ball team, and the two afore mentioned decisions are certainly different from the decision of a couple to get married. However, when we said that microeconomics attempts to provide a theory of individual decisions, above, we did not mean that microeconomics provides different theories for the decisions of bakers, sports club managers and fiancés. On the contrary, microeconomics tries to identify features of decision making that are *general* in the sense that they are common to all kinds of rational decision making, irrespective of who the decision maker is and what the decision is about. Based on this general approach microeconomics serves as the basis of many different fields of applied economics, such as environmental economics, industrial organisation or the economics of the family.

[1] Describing and explaining constitute the *positive*, evaluating the *normative* part of microeconomics. See Sect. 3.4, above.

A. Endres and V. Radke, *Economics for Environmental Studies*,
Springer Texts in Business and Economics, DOI 10.1007/978-3-642-31193-2_6,

Let us now elaborate on each of the four dimensions of microeconomic theory mentioned above.

- Microeconomics as a theory of rational individual decisions

It is an essential feature of human life that each autonomous individual takes many decisions every day. This is a burden and a privilege, at the same time. There are fundamental and trivial decisions. An example (of the former) is that most of the honourable readers of this text, at a certain point in their lives, may have decided to enrol on an environmental studies programme of a university. Instead, they could have decided to enrol on a different programme or not to go to university at all. Another example is a family deciding on whether to buy a new car instead of travelling abroad in their vacations.[2]

These decisions (and all others), have an important feature in common: the decision maker must choose between alternatives. "Alternatives" means that realising one possibility implies that the other possibility cannot be chosen. The decision in favour of a new car, for example, implies that the vacations abroad cannot be realized. (At least, this is so in the case of the family playing its part in our example.) As regards the decision to enrol in a certain university program the possibilities to do the one and leave the other are rather limited. Some people might go for two university degrees, simultaneously, but this is the maximum you can do, in most cases.

The reason for the necessity to choose and therefore being forced to decide is that the realisation of an alternative uses up resources. Once these resources are consumed for one alternative they cannot be used for the realisation of the other alternative. It is impossible "to have your cake and eat it". These resources can be of completely different kinds for different decisions. For example, the resource that forces you to decide between buying the car and going away for vacation is money. The resource that forces you to decide between different university programmes is money too, but in addition, it is time.[3]

Of course, the fact that the decision in favour of a certain alternative uses up resources wouldn't be worth mentioning if resources were abundant. It is the *scarcity* of resources, as discussed in Sect. 2.4, above, that adds a painful touch to decision making: realizing one alternative comes at the cost of not being able to realize the other alternative. The alternative foregone is interpreted to be the cost of the alternative chosen, in economics.[4] It is called "opportunity cost".

So if microeconomics is a theory of individual decision making, this is a consequence of the central theme of scarcity that we have discussed in Part I, above. Since economics focuses on the painful aspect that the decision to realize a

[2] The wish to have a new car, and to travel abroad in the vacations are "needs" in the sense of Sect. 2.1, above.

[3] These two resources are related to each other, even if the saying "time is money!" somewhat overstates the case.

[4] If there are several alternatives forgone, it is the best of these which counts as the opportunity cost.

certain alternative uses up scarce resources, it is easily understood what a criterion for "prudent decision making" is: in the economic view, a necessary condition for a decision to be prudent is that it does not use up more resources than absolutely necessary to achieve the purpose of the decision. Decisions like this do not imply any waste of scarce resources. They are termed "efficient" in economics, as has been explained in Sect. 3.1, above. Aiming to behave efficiently, people shopping for a car or deciding upon their vacations will try to find out which offers provide the best value for their money. Analogously, students try to find out which university programme will provide the best return on the invested money and time.

Of course, the idea of achieving a certain goal efficiently does not say anything about what this goal might be. Indeed, the criteria according to which a car is chosen and the relative weights of these criteria in the decision making process might be quite different for different individuals. For some people fuel efficiency might be most important. For others, size, speed, or prestige effect may be crucial. Similarly the decision to pursue environmental studies might be driven by different motives. Among these motives might be the prospects of future earnings ("Environmental experts are rich!"), personal preferences ("Environmental science is so fascinating!") or the feeling of moral obligation ("Environmental experts contribute to a better world!").

With regard to the goals individuals try to achieve, microeconomics is quite open-minded. Microeconomics presumes that individuals strive to attain as much satisfaction as possible. In the microeconomic terminology they are trying to "maximize their utility". What and how high this utility is, completely subjective. It is defined by each individual decision maker, and not by the economist who analyses the individual decision making. Microeconomics assumes that each individual agent possesses a web of tastes which enables him/her to rank alternatives according to how desirable they seem to him/her personally. This set of rankings is called "preferences". The autonomy of an individual decision maker to define his/her own preferences is called "consumer sovereignty" in the microeconomic literature even though it does not apply to consumers only, but to any decision maker.

However, there is an important exception to the rule that microeconomics is unspecific about the goal of the individual decision maker, leaving the definition of "utility" quite open. This exception applies if the decision maker under consideration is a *firm*. A firm is an economic unit combining productive factors, like labour and capital, to produce commodities and services to supply them to consumers (and other firms). In the case of firms we generally take it for granted that the utility they strive to maximize is *profit*. We will elaborate on this in Sect. 6.3, below.

We have explored some general features of microeconomics as a theory of individual decisions. We will apply these general insights in Sects. 6.2 on the consumer and 6.3 on the firm, below. Before we do that let us examine a few general observations on the second dimension of microeconomics, mentioned above:

- Microeconomics as a theory of conflict resolution and coordination

  Humans are social beings. On the one hand, each individual competes with other individuals for scarce resources. On the other hand, no individual can cope without cooperating with other individuals.

**Art work 1** David Dalla Venezia, No. 269, Oil on Canvas, 2000
This might be the illustration of three economics textbook authors (they all look alike!) struggling on whose name will appear first on the cover. Obviously, they are using brute force as a means of conflict resolution. Alternative mechanisms of conflict resolution would be the seniority principle, or – sometimes it's as simple as that – the alphabet. An alternative (and possibly more serious) interpretation would be that the painting illustrates the internal creative struggle of the artist. Interestingly, this kind of a conflict, the conflict within one and the same person, is dealt with in the arts, philosophy, as well as in poetry and prose, but not in mainstream economics. In economics, it is assumed that the decision making process of any individual does not use up emotional (or any other) resources, but is conducted "costlessly". When we talk about "conflict", in economics, we are most certainly referring to conflict between different individuals

Conflict resolution and coordination were no hot topics for *Robinson Crusoe*, at least in the time before he met *Friday*. Sparse as the resources on his island were, they were all his – at least that's what he thought, initially. However, there was no one with whom he could share the tasks of life.

Modern societies are designed in sharp contrast to the Robinson Crusoe scheme. People compete with each other for all kinds of objects of desire, like consumer products, jobs, organ transplantations, and grants to cover the costs of university environmental studies programmes. Moreover, labour is strictly divided between people who specialize in narrowly defined tasks. For example, there are environmental scientists who know most everything about the greenhouse effect but who do not know how to bake a bread, let alone how to drive a nail straight into a piece of wood. This is no problem: working on how to meet the need of society for energy without burdening the earth with excessive greenhouse gas pollution, they earn money. They can use this money to pay people who know to bake bread and to work on the nails if this is what needs to be done.

This may sound trivial but it is not. On the contrary, we are talking about a fundamental problem every society must solve. The problem consists of two elements.

The first is that society must provide a mechanism for the solution of conflicts in cases where the goals of different individuals are not compatible with each other. This is a social "must" no matter what the conflict is about: if two people want to have a unit of a certain good, it must be decided who prevails, no matter whether this unit is the last slice of bread, or the last student slot in an environmental studies programme.

The second aspect of the problem with which we are dealing here is that where ever the goals of different individuals are compatible with each other, society must provide a mechanism which makes people cooperate together like the different organs in a healthy body. This means, for example, that the decision of a consumer to buy a certain good can only be realized if this decision is complemented by firms (or other institutions) producing this good. Your decision to learn about environmental sciences and to get a qualification which makes you fit to enter the labour market for environmental specialists leads to nothing, except your frustration, if your decision is not complemented by universities' decisions to offer environmental studies programmes.

**Art work 2** Theater Bonn, Germany, FRIDA KAHLO, 2003. Director and Choreographer: Johann Kresnik (Photo: Thilo Beu). There are many wonderful photos particularly from Dance and Opera on Thilo Beu's webpage, http://www.thilo-beu.de/
In economics, we certainly acknowledge the utmost importance of coordination among individuals for their well-being. However, we take a rather prosaic look at it. The scene from the ballet shown above reminds us that there are poetic (if not romantic) alternatives. Here's the authors' understanding of utmost romance: we believe the photo shows a (somewhat farsighted) couple coordinating in jointly reading *Economics for Environmental Studies*. And don't they look happy?!

The problems we are dealing with are about how scarce goods are assigned among the people who want to lay their hands on them: is David's or Susan's application successful when the last student slot in the prestigious environmental

studies programme is assigned? The problem is also about how scarce productive resources are allocated to competing uses: does the university of Nobel City decide to stage an environmental studies programme or would it rather extend its famous medical school? A certain amount of scarce goods assigned to competing consumers (and a certain amount of scarce productive resources assigned to competing uses) is called an *allocation*, in economics.

Of course there are many possible ways to solve conflicts between individuals rivalling for scarce resources. Using the terminology that has just been introduced we might just as well say: there are many *allocation mechanisms* which may be applied. A mechanism which has a particularly long tradition in the history of human-kind is brute force. However, in the process of societies' evolution human kind became aware that different mechanisms for the regulation of conflicts have different potentials to contribute to human welfare. In this respect, brute force fared pretty badly compared to other mechanisms.

Therefore, microeconomics does not pay much attention to brute force as an allocative mechanism.

As for conflict resolution, there are many allocative mechanisms which might be considered to arrange cooperation.[5] Among these possibilities is the order of the king (of the queen!). Another mechanism is a central allocation plan designed by a national bureaucracy. Thirdly, the democratic process might decide on how to coordinate individual behaviour by majority voting. Moreover, social traditions and informal agreements may play a role in coordinating.[6] Finally, coordination might be arranged by the market mechanism.

There is not much economic analysis nowadays on monarchical rules or on central bureaucratic planning, but there certainly was, a while ago. However, there is a lot of microeconomic work on voting procedures, social traditions and informal agreements. This is done in special branches of microeconomics, called *public choice, social choice* and *institutional economics*. In the present introductory text, however, we do not pay much attention to these specialities but rather concentrate on mainstream ("bread and butter") economics. There, the focus is on the market mechanism and governmental intervention as the most important means of allocating scarce resources both in the process of conflict resolution and in the process of coordination.[7]

Let us deal with the market first (and leave government intervention to the final part of this section).

Under certain conditions, this mechanism is thought to be quite effective in managing the two social problems mentioned above, conflict resolution and

[5] Various allocation mechanisms have been briefly explained in subsection 3.5.1, above.

[6] An example of coordination by social tradition is the (admittedly somewhat old-fashioned) "ladies first" rule, coordinating the behaviour of two people of opposite gender who are about to pass through the same door. Even though it might be considered to be not completely politically correct, the rule has an impressive success rate in preventing people from bumping into each other.

[7] See subsection 3.5.2 on economic rationales for governmental intervention. In the context of environmental problems, market failure due to externalities is the most important one. These issues will be dealt with in subsection 6.5.3 and Chap. 7, below.

coordination. We will elaborate on this in Sect. 6.4, below, dealing with "the market". In the present section we make some general observations on how the market mechanism might be able to meet the need for coordination and for the solution of conflicts.

To do so we will consider an extremely simple example. Imagine that the firms in a certain market have decided to produce 50 units of a certain good, X. We won't worry about how this decision was made. This kind of a question is postponed to Sect. 6.3 on "the firm", below. On the other side of the market there are assumed to be 100 consumers and each of them would love to have exactly one unit of this product. Of course, this example is highly artificial. This might be considered to be a drawback. On the other hand, it has the advantage of being most simple. It is the simplest means of demonstrating the point at issue here: how the market mechanism can contribute to solving the two problems involved here, the one of coordination and the one of conflict resolution.[8]

In the context of this example, the coordination problem is restricting the total quantity demanded by the consumers to the total quantity of supply provided by the firms. The aspect of conflict regulation is how to decide which ones among the 100 consumers are the "lucky winners" who will receive the unit of the product in question. In a market system, the issue is not decided by a lottery but by the market price. We are looking for the level of the price for good X which cuts back the aggregate demand of the consumers to 50 units and simultaneously identifies the 50 consumers who will buy the product. This special price which equates the total quantity demanded to the total quantity supplied is called the *equilibrium price* in economics. Obviously, this price cannot be 0. As the example has been constructed, each consumer would want one unit if the product was given away for free and then the total demanded quantity would be in excess of the total quantity supplied. Excess demand is at a level of 50 units in our example if the price is 0. Let us assume the price is one monetary unit (dollar, euro, yen . . .). Then, each consumer is forced to think about it thoroughly (or fulfil the task using his/her intuition). Since the budget from which the monetary unit (let it be dollar hereafter) must be taken is limited, the question is: is a unit of good X worth the dollar? To answer this question the consumer must consider other possibilities to spend the dollar and compare the utility that this dollar generates if spent on good X with the utility this dollar provides if spent on any other good, Y.[9] On the basis of this kind of an internal evaluation process, a certain number of consumers (say, five consumers) may decide not to spend the dollar on good X but instead on an other good (or put it away into a savings account). Then, the

[8] Choosing the simplest way to make our point meets the requirement of *efficiency*, one of the most important concepts in economics, as explained in Sect. 3.1, above: we choose the way in which we achieve our didactical goal such that it confines the time our readers must devote to this issue to its minimum level. You see: we treat your time budget as a scarce resource which has to be allocated efficiently. So the rule for prudent decision making that has been explained in the previous subsection also holds for decisions in the process of economic theory building, and of textbook writing.

[9] Not being able to buy a dollar's worth of Y is the opportunity cost of buying a dollar's worth of X.

discrepancy between the quantity demanded and the quantity supplied has already somewhat narrowed compared to the situation with a price of 0, but is still at 45 units. Obviously the equilibrium price for X is higher than 1 dollar.

Assume that at a price of 43.50 dollars there are 51 consumers left who think, based on an assessment of their needs and their budgets, that good X is worth buying. Then, the equilibrium price is the lowest price at which one additional consumer can be induced to refrain from buying X. If at 44 dollars just one of the remaining 51 consumers decides to leave the market, then 44 dollars is the equilibrium price: at that price, the total quantity demanded is exactly equal to the total quantity supplied. At the same time this price draws the line according to which it is decided which of the 100 people interested in X, in the first place, actually receive it and which don't. Obviously, the people who prevail are the ones who are able to demonstrate the intensity of their needs for X by paying the equilibrium price. If all consumers had a budget of the same size, then the people with the highest intensity of need for X would be the ones who in fact received a unit of X. If different people have different sized budgets, this is not so simple. The willingness to pay depends on both, the intensity of the preference for the good and the size of the budget.

Of course, a prerequisite for the workability of the mechanism described above is that all people participating in the process respect the rules of the market mechanism as a means of coordination and conflict resolution. There must be a consensus in the society that scarce goods are allocated using the price mechanism, at least as far as goods like X are concerned. An alternative would be that goods are stolen. Another possibility is that they are distributed by the government according to a catalogue of criteria defined by a bureaucracy. A third possibility would be allocation by waiting in line. All of these procedures play a certain role in most societies. Microeconomics can be applied to analyse what's going on if any of these allocative schemes is applied. However, the focus of microeconomics is on the market mechanism (and on governmental intervention).

- Microeconomics as a theory of evaluation of resource allocation

Under the first bullet point of this introduction into the objectives and methods of microeconomics, we presented some observations on how microeconomics stylizes individual decisions: economic agents are taken to choose among alternatives, striving to achieve their goals as well as they can in light of the fact that their resources are limited. We call this behaviour *rational*, in microeconomics. What these goals are is left (almost[10]) completely down to the individual decision makers themselves.

[10] There are some basic requirements on the preferences of individual decision makers which somewhat attenuate the generality of the observation made above. However, we do not follow this line of thought here. See, e.g., Varian (2010), p. 35/36. Moreover, not all individual goals are socially accepted. There are certain constraints on individual goals defined by the law, but also by ethical principles, as well as the customs of a particular society. Notably, microeconomic analysis can also be applied to illegal behaviour. See the groundbreaking (1968) work of Economics Nobel Prize Laureate Gary Becker and, for a more recent exposition, Chaps. 11 and 12 in Cooter and Ulen (2004).

This approach is *positive*, in the sense explained in Sect. 3.4, above: it is used to describe, explain (thereby, hopefully helping to understand) and even to forecast individual behaviour. However, the approach does not say whether individual behaviour is good or bad.

An analogous statement holds for what has been said under the previous bullet point. Here, we have explained how the market mechanism might coordinate the decisions of different individuals and how it might resolve conflicts about the allocation of scarce resources among those individuals. This is also *positive analysis*; we did not say whether the market mechanism does a good job in terms of coordination and conflict resolution.

Many microeconomic analyses are confined to the positive approach but there is also an important branch of microeconomics which is *normative* (see Sect. 3.4 and subsection 3.5.2, above). Here, microeconomics tries to assess whether a certain situation is "good" or "bad". This evaluative analysis is not used very much in the area of individual decisions. Modern microeconomics does not teach people about what is good for them. The idea that individuals are entitled to decide upon their own goals, the principle of *consumer sovereignty* mentioned above, does not leave much room for microeconomic value judgements.

This is completely different where the results of social mechanisms (like the market mechanism) are concerned. There the question arises of whether the results of this mechanism in terms of conflict resolution and coordination are good or bad from the point of view of society as a whole.[11]

Of course, to assess whether a job is done well you need a criterion with which you can measure the quality of the performance. To find such a criterion is much more complicated when the quality of a situation is to be evaluated from the point of view of society as a whole than it is from the point of view of a single individual. As has been mentioned above, the issue concerning the individual decision maker is solved by microeconomics in a stunningly simple way: the decision maker is his/her own judge when it comes to assessing what is good and what is bad. However, when it comes to assessing situations from the point of view of society, we cannot get away as easily as that. The reason is that society is not a single comprehensive decision unit ("a person"), but consists of many different people. As we have observed above, each of the many members of society might follow his/her own personal goals and there may be considerable conflicts between these goals. In light of this fact, it is a difficult philosophical question when we ask how criteria might be designed according to which different situations may be assessed from the point of

[11] When we talk about "society", it must be mentioned that it consists of human beings only. Animals and all kinds of vegetation do not "count" as members of society in mainstream economics. This may be heartless but it is not as heartless as it sounds at first hearing. It does not mean that animals and vegetation are not respected and have no value from the point of view of economics. They are and they do have value, but only in so far as respect and value is attributed to them by human individuals. So animals and vegetation play a role in economics but it is a secondary role as granted by human beings in the context of an anthropocentric approach. (See Sects. 3.4 and 3.5, above, on anthropocentrism and ecocentrism.)

view of society as a whole. It is this issue which we courageously tackle in this subsection.

Compare two different allocations a certain society might be able to realize, A and B. To illustrate this, imagine that society enjoys a certain provision of consumption goods which are sold in private markets and also enjoys a certain level of environmental quality, in allocation A. In allocation B, consumption is a little lower than in A and environmental quality is somewhat higher. The question is: which allocation is better for society? In the common terminology of microeconomics you could reformulate this question to ask: is the *welfare* of society higher or lower in A than it is in B? A subsequent question is: in which situation (A, B or a third alternative C) is the welfare of society maximal? The allocation for which the welfare of society is maximal is called the "socially optimal" allocation.

Microeconomists have been working on these questions for quite a while (and very hard too). We will deal with the answers they have come up with in subsection 6.5.1, below. In this introductory section we work with the idea that good old Jeremy Bentham proposed in 1776, and which was referred to in Sect. 3.4, above. According to this somewhat cryptic but also plausible concept, an allocation is *socially optimal* if it provides the greatest happiness to the greatest number of members of the society under consideration.

This leads us to the fourth (and last) dimension of microeconomics.

- Microeconomics as a theory of public regulation

Microeconomics uses the idea of social optimality to assess the results that are produced if a society uses a certain allocation mechanism, e.g., the market mechanism. If this mechanism produced results ("equilibria") that were socially optimal, this would be a strong case in favour of the allocation mechanism under consideration leaving things as they are. On the other hand, if the equilibria produced by the allocative mechanism a society applies do not meet the criterion of social optimality, the question arises as to whether the welfare of society can be improved by governmental intervention. It is important to note, however, that "governmental intervention" is a very comprehensive term, incorporating many forms of governmental intervention. Microeconomics tries to analyse the properties of different forms of governmental intervention and to find designs that best meet the objective of maximizing social welfare.

Ooophs! This sounds awfully philosophical. Is it still economics? Yes, it is, as will become apparent when we apply these concepts in the subsequent sections.

To give you a preliminary idea of how this might work, consider environmental problems. A doctoral degree in environmental economics is not needed in order to observe that an unregulated market mechanism will not be able to secure natural resources for this and future generations. It is a safe guess that governmental intervention will be needed to protect the natural environment, and that this protection will improve social welfare. However, there are obviously different forms of governmental intervention that benefit the environment and society. Wherever the readers of this book may be located, each will know – from their home country – examples of different kinds of environmental policy instruments.

These include environmental taxes, environmental subsidies, various emissions trading programmes, requirements to apply environmentally sound technologies, and many other forms of environmental regulation. An important economic question is to evaluate these different kinds of environmental policy instruments in terms of their effectiveness and their potential to enhance social welfare. We will briefly deal with these issues in Chap. 7, below, where we discuss the microeconomics of environmental policy.

## 6.2 The Consumer

We will turn now to one of the main actors of any market economy: the consumer. The consumer is an economic agent, supplying such things as labour and capital in the market for productive factors, as well as demanding in the market for consumer goods items such as refrigerators and haircuts.[12] In discussing the economics of the consumer we apply microeconomics predominantly as a theory of rational individual decisions, in the sense explained in the preceding section. In this section we identified the scarcity of resources to be the reason for the necessity to decide among alternatives.

Let us apply this idea to the role of a consumer supplying labour. The resource which is scarce in this context is time. Each consumer must decide how to allocate his/her time budget to alternative activities all of which need time to be performed. Economically, the most crucial of these activities are work, leisure, and education. The most important benefit from work is money (as far as traditional microeconomic theory is concerned). The most important benefit from leisure is fun, hopefully.[13] The most important benefit from education is to improve the chances of making more money by working in the future.[14] Microeconomics stylizes the decisions of consumers to divide their scarce time budget optimally among the three competing uses in the sense that the total utility derived from spending the time is maximized. The situation in which a consumer spends his/her time in the utility maximizing manner is called the *equilibrium* use of time. Naturally, this equilibrium may look different for different consumers. It depends upon the relative utility a consumer

[12] Additionally, a consumer might operate in money markets, borrowing and lending. However, we do not pay very much attention to this dimension of consumer decisions. In the context of environmental issues it is not as consequential as the activities of consumers in markets for physical resources. The present textbook is designed to present economics as it is most useful to environmental studies.

[13] The German philosopher Johann Friedrich Herbart (1776–1841) said: "Boredom is the biggest sin!". This is understandable from the point of view of microeconomics because a lifetime is definitely limited. Interestingly, the economic quest for efficiency, often sneered at to be a low "purely mercantile" issue, can be interpreted as a high moral obligation.

[14] Most of our readers know that education is a lot of fun, too – isn't it? However, we do not deal with this aspect in the above brief exposition.

derives from competing allocations of time. All other things being equal, a consumer with a strong preference for consumption goods supplies more labour to the market, in equilibrium, than a consumer who does not care very much for the material world and likes best to sit meditating in the sun (weather permitting). Another determinant of the equilibrium usage of time is the wage rate. All other things being equal, more labour will be supplied (and less leisure enjoyed) the higher the wage rate.[15] If the wage for an additional hour of work is $50, a greater proportion of our cherished readership would be happy to sacrifice an hour of leisure compared to a situation with a wage rate of $5. Similarly influential for the structure of the equilibrium time budget is the extent of the wage increase that can be achieved by attaining a higher professional qualification. If a certain education is very profitable in this sense, more people will decide to invest their precious time by partaking in educational programmes where they can attain the respective qualification. Moreover, the specific time an individual may need to finish the educational programme under consideration depends on several factors, particularly the intellectual capacity of the respective individual. So the equilibrium allocation of the time budget will turn out to be different for people with different endowments of intellectual (and other[16]) resources.

Having sketched the decisions of the consumer supplying in the factor market as utility maximizing management of his/her time budget, let us now turn to the role of this economic agent demanding consumer goods. Here, the necessity to decide among competing alternatives is predominantly rooted in the scarcity of money. In order to receive a consumer good legally, one has to pay the price and, like it or not, the budget of each consumer is (more or less) limited. Resultantly, each consumer has to decide how best to divide his/her scarce budget between consuming and saving. Simultaneously, he/she must decide on which of the very many alternative commodities and services to buy as well as in which of the many alternatives he/she should save or invest. Of course, if savings already exist at a certain point of time, the consumer might decide to increase his/her budget available for consumption by reducing the stock of money saved. However, we leave savings aside in the subsequent analysis and concentrate on consumption.[17] Analogously to what has been said for supplying labour, microeconomics stylizes the decision of the consumer buying goods to be one of utility maximization. The consumer is assumed to spend his/her money optimally in the sense that he/she buys the bundle of goods for which the utility that is extracted from spending the limited budget is higher than it is for any other bundle that could be alternatively purchased with this budget. This

[15] If the wage rate gets "very high", this relationship might be turned upside down. The same holds true for "very low" wage rates. See, e.g., the section on the "backward-bending" labour supply curve in Varian (2010), pp. 174–178.

[16] If you inherit a huge fortune you probably won't worry so much about putting in a lot of work to earn money.

[17] Since saving is assumed to be zero in most introductory microeconomic analyses, the terms "budget" and "income" are used as synonyms.

utility maximizing bundle of consumption goods constitutes the *equilibrium demand* of a consumer. Of course, the structure of equilibrium demand is generally different for different consumers. This is so, because the preferences of different consumers differ. Some consumers are prepared to pay up to 200 dollars to see a soccer match, others couldn't care less. Moreover, the budgets of different consumers differ widely.

**Calculus Club: Session 1**

In this digression we address those of our readers who are familiar with the mathematical method of *calculus*. Occasionally during our exposition we call this subset of readership together to an imaginary "club session", where we use some mathematical language to make the point. So this section is more appropriately geared to people who take the environmental engineering perspective as the focus of their environmental studies, rather than to those predominantly interested in environmental law or philosophy. We hope that the members of the latter group share our assessment that the discrimination we practise hereby is a mild and tolerable one. The contents of the "calculus club sessions" are strictly supplementary (and thereby optional). The argument can be fully understood by everyone skipping the sessions. However, mainstream economics is a social science that heavily relies on the use of mathematical methods. In the main parts of this book we do without but it would be unwise to completely ignore the didactical potential of mathematical methods; thus, we address those readers who are somewhat familiar with this language.

Having said that, the members of the Calculus Club may consider a consumer buying two goods X and Y with quantities x and y. The utility he/she derives from consuming these goods is represented by a twice differentiable utility function

$$U = U(x, y).$$

According to an intuitive (even if somewhat old fashioned) interpretation, this function quantifies the satisfaction the consumer derives from consuming the goods. According to a modern (and somewhat prosaic) interpretation, U is just an index function attributing the higher numbers to the dependent variable the more the consumer under consideration likes the bundle of goods, which is represented by the independent variables. So if a consumer prefers a certain bundle $(\bar{x}, \bar{y})$ to a bundle $(\hat{x}, \hat{y})$, then

$$U(\bar{x}, \bar{y}) > U(\hat{x}, \hat{y})$$

follows.

*(continued)*

The first partial derivative of this function for any of the two variables is assumed to be positive, the second derivative for any of the two variables is assumed to be negative. The second cross derivative may be positive, negative, or zero. The first partial derivative is called *marginal utility*, in economics. Its economic interpretation is the utility generated by the consumption of a small additional amount of the product under consideration. Strictly speaking (after all, we are talking calculus here), the unit is indefinitely small.

Assume that the amount of money the consumer spends for consumption is m and the prices of the two goods are $p_X, p_Y$. Then, all the combinations of $x$,$y$ the consumer might buy are given by the equation

$$p_X x + p_Y y = m.$$

We call this equation the consumer's "budget constraint".

From all the combinations of the two goods for which the consumer is able to pay with his/her budget, he/she is assumed to choose the one providing the highest utility. This decision rule is formalized by maximizing the utility function under the budget constraint.

$$U(x, y) = \max!$$

$$s.t.$$

$$p_X x + p_Y y = m$$

The tuple of consumption quantities solving this constrained optimization problem (let us call it $(x^*, y^*)$) constitutes the equilibrium demand of the consumer under consideration.

Above we have characterized the concept of equilibrium supply (of labour) and demand (of commodities and services) of a consumer, as stylized by microeconomic theory. This has been done verbally for all of our readers and in somewhat more formal terms for the subset of our readers familiar with calculus (confined to the case of demand). We have also argued that the structure of equilibrium supply and demand might vary with the level of all kinds of determinants. In the case of equilibrium demand, the consumer's preferences and budget as well as the product prices have been mentioned.

Above, we tacitly assumed that these determinants do not change. We did not introduce any movement in terms of salary, prices of other goods, etc. This kind of an approach, describing equilibria under the assumption that the determinants of those equilibria are unchanged, is called *static analysis*.

Obviously, however, the size of each of the aforementioned determinants may change over time. That stated, the question arises as to how the consumer's equilibrium might react to changes in the determinants.

This part of microeconomic theory, analyzing how equilibria vary with varying sets of determinants is called *comparative static analysis*.[18]

Answering questions related to comparative static analysis is at the core of microeconomic theory of the consumer. However, microeconomics does not pay equal attention to each kind of aforementioned determinant.

Among the three determinants most important for equilibrium demand (preferences, product prices, and income), changes in the first receive almost no attention from mainstream microeconomics. Explaining changes in the preferences (the utility function) of consumers has been almost completely surrendered by mainstream economics to the realm of other behavioural and social sciences, such as psychology and sociology. On the other hand, the explanation of changes in prices and the consequences of such changes for equilibrium demand are central to economic analysis.[19] Changes in income and their effects on equilibrium demand play an equally considerable role in economics. In the context of the present section, we concentrate on how changes in prices affect an individual consumer's equilibrium demand.

We posit a situation in which a certain consumer achieves his/her equilibrium in consumption. Here, with given preferences, prices, and income, the consumer buys the equilibrium quantities of the available commodities and services. Unless the level of one of the determinants changes, the consumer would not want to change this situation.[20] (If we imagine that consumption takes place in different time periods – for the sake of higher plausibility –,[21] with one decision on the equilibrium demand in one period, the consumer would not want to make a different consumption plan for the next period from that which he/she has made in the previous period.)

Now assume that the price of one of the goods in the consumer's equilibrium basket increases, all other things being equal. Then, the old (pre-price change)

[18] There is a third kind of method, in addition to static and comparative static analysis. Here, microeconomics observes the trajectory along which the variable to be explained (e.g., the quantity demanded) changes in the process of moving from one equilibrium to another (or even to no other equilibrium at all). This kind of a method is called *dynamic analysis*. We do not pay very much attention to it in the microeconomic part of this book. However, dynamic analysis plays an important role in the macroeconomic part, particularly in Sects. 9.2 and 10.1.

[19] There are some (well, obviously) *extremely smart* remarks on the division of labour between economics and other social sciences in the (1993) article by Economics Nobel Prize Laureate Gary Becker.

[20] Indeed: not wanting to change behaviour given the circumstances is the "heart of the matter" in the microeconomic concept of *equilibrium*.

[21] To imagine different periods here is helpful to get the idea. However, strictly speaking, there is no *time* at all in the concept of static analysis. Acknowledging the passage of time in economic modelling belongs to the realm of dynamic analysis.

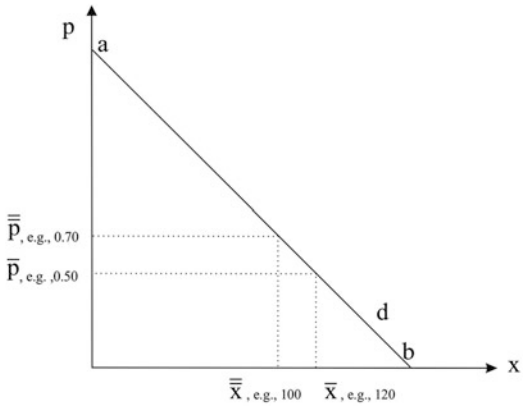


**Fig. 6.1** The individual demand curve

equilibrium cannot be upheld. The income is just not high enough to pay for the initial consumption bundle after the price of one of the elements of this bundle has gone up. It is unavoidable that the consumption level of at least one of the goods purchased by the consumer must go down, given that the price of one of the goods goes up. It is most plausible that the equilibrium quantity of the good whose price has risen actually decreases as a reaction. For example, if you imagine that the price of apples in the market increases, with all other prices remaining as they are, it is quite plausible that many consumers may revise their consumption plans: some might buy pears instead.

This inverse relationship between the price of a good and the quantity of this purchased good is often referred to as the *law of demand*. However, this is neither a law from which no deviation is possible – as is the case with natural laws – nor is it a law in the sense that deviators are punished (in the sense of criminal law). The law in the economic sense is a general observation which holds true in most cases but from which there are occasional deviations.

Of course, what has been said above for an increase in the price of a certain good can be generalized for any price change in any good.

A graphical illustration of this idea is the *demand curve* of an individual consumer, *d*. For every given price the curve indicates the corresponding quantity that the consumer under consideration demands in equilibrium. It is presented in Fig. 6.1.

For any given price observable from the ordinate of the graph, the curve shows the quantity demanded by the consumer for which this demand curve holds at the abscissa.[22] In the example, at a price of \$0.50 the consumer purchases 120 units of the good *X* in equilibrium. In case of a price increase to \$0.70 the quantity demanded drops to 100 units.

[22] The abscissa is the horizontal, the ordinate is the vertical axis.

Putting it in more general terms you might say that at a price of $\bar{p}$ the equilibrium quantity demanded by the consumer under consideration is $\bar{x}$. If the price increases to $\bar{\bar{p}}$, equilibrium demanded quantity drops to $\bar{\bar{x}}$. The law of demand is respected in that $\bar{\bar{p}} > \bar{p}$ and $\bar{\bar{x}} < \bar{x}$ hold.

The demand curve is the graphical representation of the *demand function*, $x = d(p)$. The equation for the demand function is illustrated in Fig. 6.1 as $x = 170 - 100p$. From the way we wrote this equation and the way we interpreted the demand curve, we note $p$ is the independent variable and $x$ the dependent. Considering this, the illustration of the demand function in Fig. 6.1 is somewhat unusual, because it would be more conventional to plot the independent variable on the abscissa, and the dependent on the ordinate. Well, economists sometimes make an exception. The exposition presented here is very traditional, and economists have gotten so used to it that they do not even realize anymore that it is unconventional. As long as the curve is monotonically decreasing, however, it doesn't really matter which way you position the independent and the dependent variable.[23] Indeed we will use an interpretation below where the quantity is understood to be the independent variable of the demand function and the price to be the dependent.

In our example, the demand curve is assumed to be linear. This is done solely for convenience. The demand curve can take any shape: all that is required is that it is downward sloping, obeying the "law of demand", as was discussed above.

There are two points on the demand curve that are immediately eye-catching: these are the points at which the curve intersects the abscissa and the ordinate, respectively. Each of the two has its own economic interpretation which might be worth noting.

Consider the intersection of the demand curve with the ordinate first, $a$. If the price is at a level of $a$, the quantity demanded is 0 (and stays at 0 for any price higher than $a$). Since $a$ is the lowest price that prevents the consumer from buying $X$, $a$ is called the "prohibitive price".

At the other end of the demand curve you find the quantity, $b$, where $d$ intersects the abscissa. $b$ is the quantity the consumer demands if $X$ is given away for free. (In order to make the consumer ask for more, you would have to pay him/her!) Quite descriptively, $b$ is called the "satiation quantity".

Interpreting the demand curve, you should always keep in mind that illustrating how the quantity demanded of a certain good depends upon the price of this good does not say that this price is the only determinant of this quantity. Instead, all other determinants are assumed to be constant in the process of interpreting the demand curve.[24]

[23] Please remember that the property of being monotonically decreasing is assured by the *law of demand*.

[24] Academic economists sometimes love to show off their high level of education. A well established trick to make an impression is to occasionally intersperse some Latin terminology. So instead of saying "all other determinants assumed to be unchanged", you might say "*ceteris paribus*".

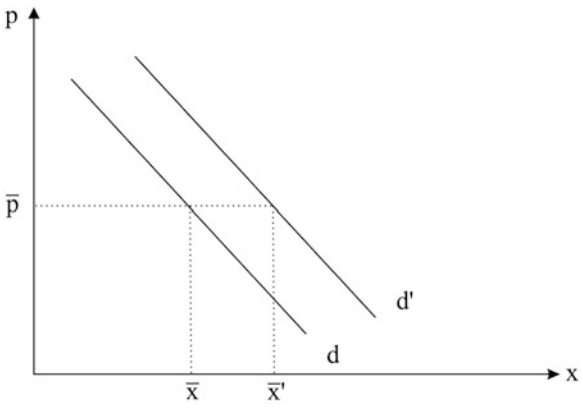


**Fig. 6.2** Comparative static analysis of demand as income increases

When the demand curve is written as $x = d(p)$, this is shorthand for the values of all the variables affecting the demanded quantity of $X$ beyond the price ($p$) of $X$ being assumed to be constant. These variables are not explicitly mentioned in the equation. However, they are nonetheless "silently kept in the economist's mind".

Among these other determinants are the prices of all other goods and the income of the consumer. Importantly, the preferences of the consumer are assumed to be constant but this goes without saying in mainstream economics.

According to what has been said above, we read the reaction of a consumer in terms of his/her demanded quantity for a certain good along the demand curve if the price of this good changes. We "move along a given demand curve", observing different points along its course.

However, what happens if one of the other determinants of the demand for the good under consideration (other than the price of this good) does not sit still, as we have assumed in defining the demand curve? Then, the demand curve shifts. To see this, we apply some comparative static analysis, as referred to above. Consider a change in income. Normally, we expect that for every predetermined price of $X$, the quantity demanded, $x$, by a certain consumer increases when his/her income, $m$, increases. If $X$ is a "normal" good in this sense, the demand curve shifts "to the north-east", i.e., from $d(p)$ to $d'(p)$ if the income of the consumer whose demand curve we analyse increases from $m$ to $m'$.[25] This kind of a shift is illustrated in Fig. 6.2.

It is not important whether the shift shown in the figure is a parallel one. All that matters in the present context is that the new demand curve $d'(p)$ indicates a higher equilibrium demanded quantity for every price, compared to the initial curve $d(p)$.

[25] A positive relationship between the quantity demanded at a predetermined price and income is plausible but might not hold for every good. Think of the market for used shoes. Goods for which demand goes down as income goes up are called *inferior goods* in economics.

In Fig. 6.2, for any price of $\bar{p}$, the demand is at $\bar{x}$ given income is at $m$ and at $\bar{x}'$ if income is at $m'$. $m' > m$ and $\bar{x}' > \bar{x}$ hold.

Above, we have presented a little exercise in comparative static analysis: one of the determinants assumed as being constant in the static analysis of the demand curve has been allowed to vary. This determinant is income.

Of course, analogous comparative static analyses can be applied with respect to all other determinants of the quantity demanded which are assumed to be constant in the consideration of a demand curve.

An issue very often discussed in economics is what happens to the demand curve of a certain good $X$, if the price of another good, $Y$ changes. Depending on how the two goods are related to each other in terms of their suitability for satisfying the consumer's needs, an increase in the price of $Y$ may stimulate or attenuate the demand for $X$. In the former case the demand curve of $X$ would shift outward (to the north-east), whereas in the latter case it would shift inward (to the south-west). A case where the demand for $X$ increases as the price of $Y$ increases might be observed if $X$ and $Y$ are two different kinds of TV sets with comparable quality. Then, $X$ and $Y$ are said to be *substitutes* in microeconomic terminology. A case where the demand for $X$ decreases if the price of $Y$ goes up might be observed if $X$and $Y$ are hardware and software. (An old-fashioned example is bread (popcorn!) and butter.) In this case, $X$ and $Y$ are called, in microeconomics, *complements*.[26]

For now, we will take a hiatus from comparative static analysis and return presently to the static analysis of the demand curve, taking a look at it from a different perspective.

In the interpretation of the demand curve, presented above, we have always taken the price to be the independent variable and the equilibrium quantity demanded to be the dependent variable: given a certain price the consumer wants to buy a certain quantity. So the curve was read "from the ordinate to abscissa". This is appropriate in the present context, explaining the decisions of a consumer as an important agent in the market.

However, it is also possible to interpret the demand curve by reading it "from the abscissa to the ordinate". This is a nice little exercise in mental flexibility. Moreover, it will turn out to be useful in Sect. 6.5. There, we proceed from describing what is going on in the market (positive analysis, see Sect. 3.4, above) to its evaluation (normative analysis, as is also dealt with in the section referred to above).

Consider the point on the demand curve shown in Fig. 6.1, as has been interpreted above, with the price being 0.50 and the quantity 120. (In more general terms the price/quantity combination is $(\bar{p}, \bar{x})$.)

[26] Our honourable readers are invited to graphically illustrate the comparative static analysis of demand for $X$ as the price of $Y$ increases. They might proceed analogously to what we have done in Fig. 6.2 referring to an increase of income. They might also distinguish the case of substitutes from the case of complements.

How about letting price and quantity switch the roles as independent and dependent variables? Then, we arrive at the following interpretation: if the consumer has bought a quantity of 120 ($\bar{x}$), then the price he/she has been willing to pay *for the last unit* of $X$ is 0.50 ($\bar{p}$). Why? Try to prove the opposite: to do that, claim first that the willingness to pay for the last unit is lower than 0.50 ($\bar{p}$). Wrong! If it were true that the consumer was not willing to pay 0.50 ($\bar{p}$) for the last unit, then he/she would not have bought it. After all, we assume that the consumer strives for utility maximization. Second, what if one claims that the willingness to pay for the last unit is higher than 0.50 ($\bar{p}$), at a quantity of 120 ($\bar{x}$). Wrong! If the willingness to pay were higher than 0.50 ($\bar{p}$) at a quantity of 120 ($\bar{x}$), then a consumer would still buy this last unit at a price which is a little higher than 0.50 ($\bar{p}$). However, he/she doesn't, which can be seen in Fig. 6.1. If you observe a price which is a little higher than 0.50 ($\bar{p}$), then equilibrium demanded quantity drops below 120 ($\bar{x}$). It can't be otherwise because the consumer operates under the law of demand.

Of course, what we have argued for one point on the demand curve (0.50, 120), ($\bar{p}$,$\bar{x}$) can be just as easily argued for any point on that curve. For each predetermined quantity you can read the consumer's willingness to pay for the last unit consumed from the demand curve. In the usual microeconomic terminology (inspired by calculus) we call the last unit *the marginal unit*. So the demand curve can be interpreted as the *marginal willingness to pay*-curve. It must be mentioned that our little mental exercise of exchanging the roles of $p$ and $x$ as the independent and dependent variables, respectively, was made possible courtesy of the law of demand: if the curves were not running monotonically you could not transform the unique mapping of prices into quantities into a unique mapping of quantities into prices.[27]

For the benefit of terminological clarity economists do not shy away from any effort and thereby created a special word for the demand curve as interpreted above: if we read the demand curve such that the quantity is the independent and the price the dependent variable we speak of the *inverse demand curve*. Take the example we used above. If $x = 170 - 100p$ is the initial demand curve,[28] then $p = 1.7 - x/100$ is the inverse demand curve. In general terms, if the direct demand curve is $x = d(p)$, then the inverse demand curve is $p = d^{-1}(x)$. Very often, however, the "$-1$" in the exponent of the symbol of the function, "$d$", is not written. There, the authors trust that the readers will understand from the context, whether it is the direct demand curve or the inverse demand curve that is referred to. We partake in this tradition of "expository sloppiness" in the notation of graphs. There, a curve labelled "$d$" might sometimes be interpreted as being the direct demand curve, and sometimes to be the inverse demand curve. We do that in order to economize in

[27] To see why, imagine that the demand curve is U-shaped instead of being monotonically downward sloping. After completing this exercise, you might immediately forget about u-shaped demand curves; they violate the law of demand!

[28] The initial demand curve (the one where $p$ is the independent and $x$ the dependent variable) is sometimes called the *direct demand curve*.

terms of the number of graphs being used and believe that the likelihood of thereby creating misunderstandings is negligible.

## 6.3 The Firm

Having dealt with the consumer as the first main actor in the drama of microeconomics we are now going to turn to the second, the firm (there is no one to stop these dedicated authors and their enthusiastic readers!). The firm is an economic decision maker buying in the markets for productive factors, labour, and capital, and selling in the markets for commodities and services. Accordingly, the firm is the counterpart of the consumer who has been defined to sell in the first and to buy in the second market.

It is worth noting that markets for commodities cannot always be distinguished from markets for productive factors, by looking at the physical properties of the product under consideration. For example, electricity is certainly bought by firms as well as by consumers. So if electricity is demanded by a household then the market for electricity is perceived as a market for a consumer good. On the other hand, if electricity is demanded by the firm, electricity is perceived as being bought in the market for productive factors. Resultantly, the same entity can be interpreted from different perspectives and it is predominantly the context of the discussion (of the economic model under consideration) that matters in terms of terminology.

In discussing the role of the consumer we treated his/her supply in the market for productive factors only in passing and spent most of our efforts on explaining the consumer buying in the market for commodities and services. For the firm we follow a complementary pattern. For simplicity, we take it to the extreme, completely ignoring the role of the firm in the market for productive factors. We concentrate exclusively on how the firm acts in supplying commodities and services to the market. Just as we have done with regard to the consumer, we use microeconomics as a theory of individual rational decisions dealing with the firm.

So what does a firm have to decide, operating as a supplier in the market for commodities and services? Obviously, in order to be able to supply, the firm has to produce. So the decisions of the firm are related to the questions of *what*, *how much,* and *how* to produce. According to what has been said about the objective of the firm, in Sect. 6.1 above, the firm strives to find the answer to these three fundamental questions that maximizes the firm's profit. The firm's profit is defined to be the difference between the revenue the firm receives from selling its products in the market, and the costs that are generated by the fact that the firm has to pay for the inputs used to produce the firm's output.[29] (The terms "output" and "production" are used as synonyms in microeconomics.)

[29] In reality you might not be exactly sure what these revenues and these costs are going to be. In this introductory exposition, we ignore problems of uncertainty, to keep things simple.

For simplicity, we do not deal with the decision of the firm on what to produce.[30] We focus on the decisions of a firm that has already decided to produce a product, $X$. Regarding the question as to how to produce, we can be quite brief in the present context. Generally there will be more than just one method of production for the good under consideration. If you compare how a certain product, e.g., steel, is produced at one point of time in different countries, or at different points of time in the same country you will realize that different productive methods are used. These productive processes are particularly distinguished from each other according to how much labour and capital is used to produce one unit of the product (say, one ton of steel) and what kinds of labour and capital are used. Typically, you find more capital and less labour being used if you compare recent production methods to earlier ones. The same tendency is revealed if you compare production in industrialized countries to production in developing countries. So a firm that plans to produce $X$ must choose which of the available production processes to use. The decision rule is quite clear: it follows from the profit maximization objective of the firm. Since costs reduce profits by definition, the firm selects the productive process with which it is able to produce at minimum cost. *Cost minimization* is a prerequisite of profit maximization. Which one of the many available production processes qualifies to be cost minimal depends on the prices the firm has to pay for the various inputs used and on the *productivity* of these inputs within the context of a specific production process.[31]

Cost minimization as a rule governing how to produce holds independent of how much is produced. Accordingly, our statement that each firm chooses to apply the cost minimizing mix of available production processes is true, independent of what level of output the firm decides to produce. Putting this into mathematical terms we can derive the *total production cost function*. As the dependent variable, this function indicates the minimum amount of cost, $C$, which is incurred for any level of production, $x$, as the independent variable. The total cost function is written as $C = C(x)$. By specifying this function for (small) additional units of production we arrive at the *marginal production cost function*, $MC(x)$.[32]

Since the cost function is defined for any level of output you cannot answer the question what the *profit maximizing* quantity, $x^*$, might be, using the cost function

[30] Also, the terms "production" and "supply" are used as synonyms. This is so, because we ignore that the firm may put part of its production into storage (or may throw it away). Everything produced is assumed to be brought to the market. Moreover, nothing brought to the market is taken out of storage.

[31] The *average* productivity of an input is the amount of output produced per unit of this input. If the level of output is denoted $x$ and the level of labour is denoted $l$, then the labour productivity is $x/l$. A related concept is the *marginal* productivity of an input. It indicates by how much output increases as the quantity of this input used in the production process is increased by one (small) unit. *Calculus Club mini session*: marginal productivity of labour is $\partial x/\partial l$.

[32] *Calculus Club mini session:* the marginal cost function is the first order derivative of the total cost function.

only. However, using this function definitely contributes to the solution of the puzzle.

Stay tuned:

Put yourself into the shoes of the owner of a firm. Before production actually starts you sit down and consider the options. Applying the method established in microeconomics to identify the profit maximizing quantity of production, you take the "piecemeal" approach. You conduct a little "thought experiment" by considering an arbitrarily chosen level of production, $\bar{x}$, and you ask the following question: if I produced $\bar{x}$, would the last unit produced make a positive contribution to my profit? To answer this question, you compare the revenue this last unit generates with the cost that has to be spent for the production of this unit.[33] The revenue generated by the last unit is called *marginal revenue,* while the cost generated by this unit is called *marginal cost*. The difference between the two is *marginal profit*.

Consider the case where the marginal revenue is higher than marginal cost; in this case, the last unit under consideration obviously contributes to the profit of the firm, meaning marginal profit is positive. Then, it is worthwhile to increase production by one further unit. If this also provides positive marginal profit then further expansion of your production plan is warranted.

If, on the other hand, the marginal profit of the production unit under consideration is negative you will refrain from producing it, assuming you strive for profit maximization. So according to this reasoning a positive marginal profit is a signal to the profit maximizing firm to expand the quantity it plans to produce, whereas a negative marginal profit is a signal to reduce planned output. Accordingly, the profit maximizing quantity is defined by zero marginal profit. For the last unit supplied marginal revenue equals marginal cost. Here, the firm does not have any incentive to change its plan (i.e., the firm is in equilibrium). As the owner of the (albeit imaginary) firm, you are "there", once you have identified the level of production which meets the equilibrium condition "marginal revenue = marginal cost". It has been a long day then, but it was definitely worth it!

We briefly explained the concept of marginal cost, above. In order to operationalize the equilibrium condition "marginal revenue = marginal cost", we must be somewhat more specific with regard to revenue. Revenue is defined to be the product of price and quantity of the good sold. In the simplest setting, these two variables determining revenue are independent from each other. Here, the price for the product is determined in the market, as has been sketched in the previous section and as will be elaborated in the coming section. The firm can supply any quantity of the product at this ruling price. However, it cannot affect the price in any way. Instead, it has to take it as it is. Therefore, the firm is often said to be a *price taker*. Specifically, the price is independent on how much the individual firm supplies to the market. Moreover, the market price is the same for all firms within the industry producing the good under consideration, $X$. This setting is quite plausible if the firm

[33] Again, we ignore uncertainty.

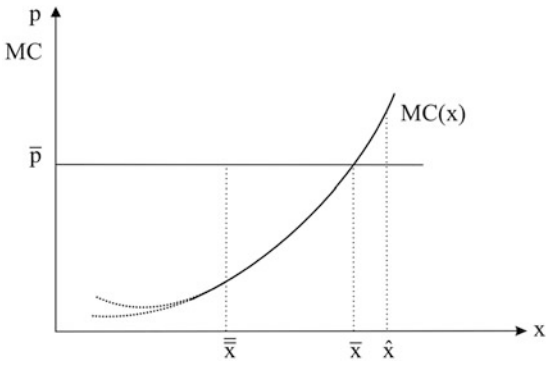


**Fig. 6.3** Equilibrium output decision of a perfectly competitive firm

under consideration is very small compared to the totality of firms producing the same good. Consider a single farmer supplying to the world coffee market. You might also think of worldwide market places on the internet. Price comparison sites and online selling portals accumulate offers from very large numbers of suppliers for many different products. Regardless where the merchant is located, he/she is able to sell to customers throughout the world. Therefore, there are hundreds and thousands of companies offering the same product, meaning these firms have to accept the lowest price on the platform used as a reference price for their own price. A market structure which is characterized by this feature, where an individual producer is "a drop in the ocean", is called *perfect competition*.[34] If a firm operates in a perfectly competitive market, the marginal revenue is identical to the market price: selling an additional unit of output generates revenue which is exactly equal to price.[35]

Other kinds of market structures considered in microeconomics are *monopoly* and *oligopoly*. In the former market structure the firm under consideration is the only one supplying in the market. Consequently, it can exercise considerable influence over the market price. The latter market structure catches the case where a few firms operate in a market where each of them has certain influence on the market price and each of them must be on the guard, observing what the other firms do. However, we will concentrate on the simplest case of perfect competition and ignore monopoly and oligopoly.

Using the assumption that the market price is independent of the quantity supplied by the firm under consideration, we can illustrate graphically what we have said on the nature of profit maximizing output (Fig. 6.3).

[34] Within the context of the present section ("the firm"), we have emphasized above that firms are price takers in a perfectly competitive market. For the sake of the completeness of the definition of a perfectly competitive market, let us add that consumers are also assumed to be price takers in this market structure.

[35] Closer inspection (from which we refrain here) reveals that this statement does not hold for other market structures than perfect competition. See, e.g., Varian (2010), pp. 439–445, 497–506.

In this figure, the price $\bar{p}$ is drawn as a horizontal line as a consequence of the fact that it does not depend on individual firm output, $x$, which is measured on the abscissa. The curve $MC(x)$ illustrates the marginal cost of the examined firm producing $x$. The dashed parts of the curve indicate two alternative shapes of the curve. They illustrate that it is not essential for the argument explained above how the curves run "far away" from the point where the marginal cost curve intersects with the price line.[36] At this intersection, the equilibrium condition "price = marginal cost" is met. The quantity for which the equality holds is called $\bar{x}$ in the figure. So, $\bar{x}$ is the answer to the question of how much a profit maximizing firm will produce. This level is contingent on the price for $X$ being $\bar{p}$, a market structure of perfect competition (implied by the fact that $\bar{p}$ is constant), and by the given values of the input prices and of technology (implied by the shape of the marginal cost curve).

The figure can be used to graphically illustrate the "piecemeal approach", so as to identify the profit maximizing output level (which has been verbally explained above). Take the production quantity $\underline{x}$ as the first example. At this quantity level, price exceeds marginal cost, i.e., marginal profit at $\underline{x}$ is positive. Flowing from this, production is too low at $\underline{x}$, i.e., the production plan has to be extended to serve the goal of profit maximization. On the other hand, at production level $\hat{x}$ marginal profit is negative. As a result, the production plan has to be contracted in order to move into the direction of the profit maximum. Obviously, for this kind of an argument to be valid the marginal cost curve must be upward sloping in the vicinity of its intersection with the price line.[37]

[36] This is an example of the fact that the logical relationships dealt with in microeconomics are often more general than suggested by the graphical illustrations used to visualize these relationships. Another example was mentioned when we discussed the comparative statics of demand as income increases, using Fig. 6.2. There, we mentioned that the demand curves have been drawn to be parallels for simplicity only. The effects of changes in income on the quantity demanded at a given price also hold for (downward sloping) demand curves not parallel to each other.

[37] Not so obvious? To solve the puzzle, you might draw a figure with $\bar{p}$ as the horizontal price line and add a *downward sloping* marginal cost curve. Then, denote the quantity for which your marginal cost curve intersects the price line as $\bar{x}$, as we did in Fig. 6.3, above. Equipped with these analytical instruments you might find out what happens to the firm's profit if it increases output a little, going from $\bar{x}$ to $\hat{x}$, or decreases output a little, going from $\bar{x}$ to $\underline{x}$. You will see that profit goes up for both deviations from $\bar{x}$. *Calculus club mini session:* If marginal cost decreases in the region of its intersection with the price line the second order condition for a profit maximum which will be derived in the subsequent calculus club is violated. In fact, the firm's profit is at a *minimum* at $\bar{x}$.

**Calculus Club: Session 2**

Just as we did in the previous section for the consumer we can summarize what we have just said about the firm in somewhat more formal terms:

The profit of the $X$-producing firm, $\Pi$, is defined to be revenue, $R$, minus cost, $C$. Revenue and cost are both written to be functions of output, $R(x)$ and $C(x)$. Revenue is defined to be price times quantity sold, i.e., $R(x) = px$ holds. Since we consider the market structure of *perfect competition,* the price cannot be affected by the individual firm, such that $p$ is a constant. Given that the profit maximization problem of the firm can be written as

$$\Pi = px - C(x) = \max!$$

The first order condition is

$$\partial\Pi/\partial x = p - \frac{\partial C}{\partial x}(x) = 0.$$

Since the derivative of the cost function is the marginal cost function, MC the first order condition can be written as

$$p = MC.$$

The solution to this first order condition is $x^*$.

The second order condition is

$$\partial^2\Pi/\partial x^2 < 0,$$

i.e.,

$$\partial MC/\partial x > 0.$$

The first order condition given above is the formal expression for the intersection of the marginal cost curve and the price line in the figure. The second order condition is the formal expression for the requirement that the marginal cost curve is upward sloping in the vicinity of the solution.

We have sketched the output decision of a profit maximizing firm. In the process of this analysis the determinants of the profit maximizing level of production have been assumed to be constant. These determinants are the price of the output as well as the marginal cost curve. In turn, the determinants of the latter, input prices and technology, have also been assumed to be constant.

Just as we have done through the analysis of the consumer in the previous section we can now analyse how the firm adjusts its output decision if these determinants

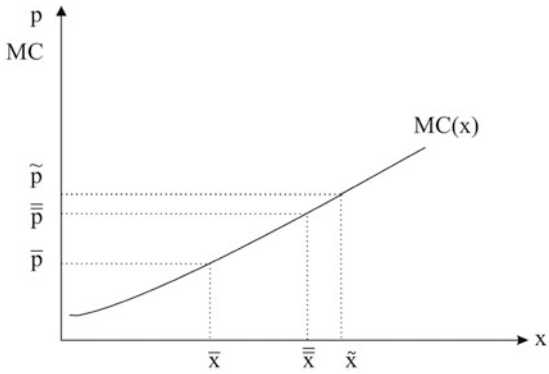


**Fig. 6.4** Equilibrium supply for alternative prices

change. We focus on how changes in the output price affect a firm's equilibrium supply. Therefore, we choose the counterpart to the analysis of a consumer where we analyse how changes in the product price affect the decisions of the consumer in terms of demand.

The rules according to which the firm chooses the quantity supplied do not depend upon the level of output price p. However, the quantity that meets the equilibrium condition "price = marginal cost" changes if the price changes. This can easily be seen in Fig. 6.4 where the equilibrium supply of the firm is shown for alternative prices, $\bar{p}, \bar{\bar{p}}, \tilde{p}$. For each of the alternative and predetermined prices the quantity supplied can be read off the (increasing part of the) firm's marginal cost curve.

Dealing with the consumer in the previous section we defined the *demand curve* to be the curve indicating the quantity demanded by the consumer at each price for the product under consideration. Analogously, the *supply curve* is the curve indicating the quantity supplied by the firm under consideration at alternative product prices. In general terms, this is written as $x = s(p)$ and called the *supply function*. An example is $x = 2p$.

It follows from the brief analysis given above that the supply curve is generated by the increasing part of the marginal cost curve of the firm.

We say "is generated by" instead of "is identical to" for the following reason: in our definition of the supply function the price is the independent and the quantity the dependent variable. In the marginal cost curve, the quantity is the independent variable and the marginal cost the dependent. So the marginal cost function, $MC = MC(x)$, is not identical to the direct supply function, $x = s(p)$ but identical to the *inverse supply function*, $p = s^{-1}(x)$.[38] An example is the marginal cost

[38] In the previous chapter we noted that the monotonicity of the demand curve is a prerequisite for the transformation of the direct demand function into the inverse demand function. Analogously, going from the direct supply function to the inverse supply function requires monotony. This requirement is met since for the constitution of the supply curve we use the monotonically increasing part of the marginal cost curve only.

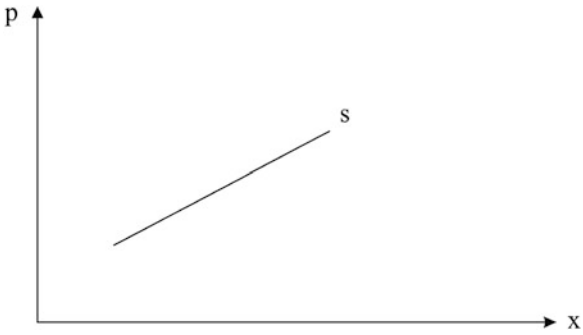


**Fig. 6.5** The supply curve of the firm (at one time, this is an exercise in economics just as it is an exercise in minimal art)

function $MC = x/2$.[39] The corresponding inverse supply curve is $p = x/2$. This function is the inverse of the direct supply function mentioned in the example given above, $x = 2p$.

In analogy to what we have said in the section about the consumer, we use the symbol "s" for supply in the graphical illustration not distinguishing between the direct supply function, $s(p)$, and the inverse supply function, $s^{-1}(x)$. The curve illustrates the relationship between price and quantity in the context of supply and it will be always clear which of the two variables is taken to be the independent and which to be the dependent.

As a result, we redraw the relevant part of the marginal cost curve and label it $s$ for supply.[40] If you read the curve "from the ordinate to the abscissa", working with $p$ as the independent variable you are dealing with the direct supply curve. If you read "from the abscissa to the ordinate", instead, using $x$ as the independent variable you are considering the inverse supply curve (Fig. 6.5).

Above we have taken the quantity of a certain good $X$ supplied by an individual firm to be determined exclusively by the price of this good. The relationship between these two variables has been graphically illustrated by the supply curve. Of course, this mapping of prices to equilibrium quantities for the individual firm is analogous to the price-quantity mapping for the consumer elaborated in the discussion of the demand curve. This analogy is no big surprise because, obviously, the consumer demanding $X$ and the firm supplying $X$ have to be (and will be!) brought together in the market. We deal with this in the next section.

[39] This is a particularly nice one because it is monotonically increasing. Sometimes U-shaped marginal cost curves are used. An example is $MC = x^2 - 10x + 30$.

[40] The generality of this statement is attenuated by the fact that maximal profit might be negative for the firm if the price is too low. At these prices the equilibrium supply is 0 and the marginal cost curve of the firm, even if increasing, is not part of the supply curve in this area. These considerations are relevant for U-shaped marginal cost curves, as the one mentioned in footnote 27, and are further differentiated if you distinguish between "long run" and "short run" analysis. However, we do not deal with these issues here and refer instead to the literature. See, e.g., Varian (2010), pp. 398–402.

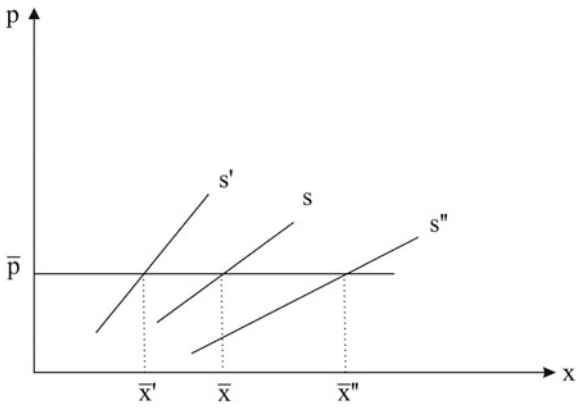


**Fig. 6.6** Comparative static analysis of the individual supply curve

Before we do that let us briefly observe an additional analogy in the economics of the consumer and the firm. If we recall our discussion on the consumer, we will note that showing the demanded quantity depending on the price of the good under consideration does not say that the price is the only determinant for the equilibrium quantity demanded. This applies equally to the firm: introducing the supply curve as an analytical tool showing how the price of a good determines the quantity supplied in equilibrium by a firm does not suggest that the quality supplied is determined by the price of the good under consideration exclusively. Writing $x = s(p)$ is just a shorthand where the determinants of $x$, other than $p$, are assumed to be constant in the course of the analysis and therefore are not made explicit. Analogously to what has been said explaining the demand curve in the previous section, however, these other determinants are silently treasured in the economist's mind.

Let us now briefly deal with these other determinants of the quantity supplied.

Obviously, how much a firm will be willing to supply in terms of $X$ depends on the cost of producing $X$. This intuition is wonderfully compatible with our earlier observation that the supply curve mirrors the marginal cost of production. However, the cost of production is not invariably determined by the laws of nature. On the contrary, it depends on factors which are subject to human decisions and might thereby change. An important issue in this case is the prices for the inputs that the firm uses in the production process. All other things being equal it is clear that production cost will increase if the prices for labour and/or capital increase.[41] If this happens, both the marginal cost and supply curves shift upwards. This is illustrated in Fig. 6.6 by the shift of the supply curve from $s(p)$ to $s'(p)$.

[41] Microeconomics contributes to a better world by stimulating the dialogue among different generations: ask your grandparents what happened during the world wide "oil crisis" in the year 1973. (A contemporary worry is what might happen to the cost of all kinds of electronic devices if the prices for rare metals go up.)

Another determinant of production cost, in addition to input prices, is technology. Technology changes over time as a result of research and development activities. Technical progress may take many forms. One of these forms induces production processes to economize on the inputs used for production. If the cost of production decreases due to a shift to modern technology, the marginal cost curve shifts downward, which in turn results in the same outcome for the supply curve. In the graphic above, we illustrate the supply curve for the advanced technology by $s''(p)$. (Input prices are supposed to be at what they were as "embedded" in the initial supply curve, $s(p)$.) You can read from the graph the impact changes in the determinants of production cost will have on a firm, with respect to its willingness to supply at a predetermined price. Given the price is at $\bar{p}$ the firm produces a quantity of $\bar{x}$ in equilibrium in the situation where input prices and technology are at their "old" levels. Given that input prices are at the "new" (higher) levels but technology is as it used to be, the equilibrium quantity supplied by the firm at $\bar{p}$ goes down from $\bar{x}$ to $\bar{x}'$. On the other hand, if input prices stay at their "old" levels and cost saving technical progress is introduced, then equilibrium quantity supplied at $\bar{p}$ goes up from $\bar{x}$ to $\bar{x}''$.[42]

## 6.4 The Market

> As a child I'd asked my Mom why was this so? Why the trains didn't stop in Mt. Ephraim any longer?
>
> Mom laughed. "Oh, ask me! As if I'd know."
>
> Then, for Mom always pondered our questions to her, even those she couldn't answer: "I think it has to do with the economy, Nikki. 'Supply and demand.' You can ask Dad, he will know."
>
> I was reluctant to ask Dad such questions. He'd squint at me suspiciously as if, at school, I'd already learned the answer and was testing him. Or, worse, he'd provide such a long and complicated answer I couldn't make sense of it. "Supply and demand" was what it all boiled down.
>
> Joyce Carol Oates, Missing Mom, New York (Ecco/HarperCollins Publishers), 2005, p. 256

Above, we have dealt with an individual consumer and an individual firm, respectively. The market we are about to discuss is an institution that arranges the coordination of decisions of these two kinds of agents, as has been explained in Sect. 6.1. Generally, however, markets do not coordinate one firm with one consumer but many firms with many consumers. This is particularly so in case of perfectly competitive markets, which is the specific kind of market structure to which this analysis has been chosen to be confined. So in order to explain the

[42] At this general level of discussion it is not clear what happens if input prices increase and technology improves, simultaneously. Then, the "new" supply curve will shift to a position somewhere between $s'$, and $s''$. Whether, compared to the "old" equilibrium quantity, $\bar{x}$, the "new" quantity increases or decreases depends upon which of the two countervailing effects prevails.

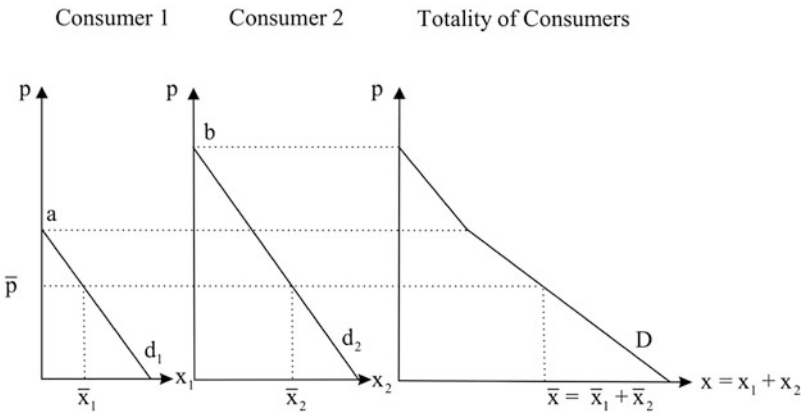


**Fig. 6.7** Deriving the market demand curve by horizontal aggregation of the individual demand curves

market mechanism we must first extend the analysis of the consumer as well as the analysis of the firm from one decision maker to many.

Achieving that goal is facilitated by the fact that the analysis given for the consumer and the firm, above, was not meant to apply to one specific individual only. On the contrary, it was meant to be a general statement that respectively applies to all consumers and to all firms. Therefore, the idea of how the totality of the consumers behave in the market is generated by the idea that all the individual consumers behave as the "typical" consumer scrutinized above. The same idea analogously holds for the firm.

Let us be specific on this analytical step from one individual agent to the totality of agents of this kind and firstly address individual consumers. For those, we have explained how equilibrium demand is determined. Specifically, it was shown how the quantity of a product demanded by *an individual consumer* depends upon the price of this product, all other determinants assumed to be constant. The graphical illustration of this economic concept has been introduced as the *demand curve*. So if we want to explain how the demanded quantities of all consumers of $X$ taken together depends on the price of $X$ we somehow have to "add up" all the individual demand curves to arrive at the market demand curve. The market demand curve would then indicate how the quantity of a product demanded by *the totality of consumers* depends on the price of this product, all other determinants assumed to be constant. The kind of "adding up" necessary to go from the individual demand curve to the market ("total") demand curve, is done by a procedure called *horizontal aggregation,* as explained below.

For graphical simplicity we take the case where there are only two consumers constituting the demand side of the market, even though the argument is meant to apply to the case of many consumers (Fig. 6.7).

In the section on "the consumer" (6.2), we dealt with the decision of a single agent. The demand curve of this individual was labelled "*d*". Now that we extend the analysis to cover two consumers we need a somewhat extended notation. We

denote the demand curve of the first decision maker "$d_1$"and of the second decision maker "$d_2$". If we make a statement which refers to the demand curve of either of the two consumers, we write "$d_i$". The "$i$" is a general expression which might take on the form of one, or of two.[43] To distinguish the *individual* demand curve from the preceding section, dealing with one consumer only, "$d$", from the *total* demand curve in this section we write the latter with a capital, "$D$".

To understand the procedure of horizontal aggregation mentioned above, take any arbitrarily chosen price, $\bar{p}$, as an example. Given this price the demanded quantity of consumer 1 is $\bar{x}_1$ and the demanded quantity of consumer 2 is $\bar{x}_2$. You can read these individual demanded quantities from the individual demand curves of the two consumers, $d_1(p)$ and $d_2(p)$, respectively. So in order to arrive at the total demanded quantity, given the price $\bar{p}$, the individual demanded quantities must be added up. Total demand at $\bar{p}$ is $\bar{x}$ with $\bar{x} = \bar{x}_1 + \bar{x}_2$. Consequently, a point with the coordinates $(\bar{p}, \bar{x})$ is a point lying on the total demand curve, $D(p)$, for the product $X$.

Of course, the logic of this procedure in attaining information on the total quantity demanded does not depend on the level of the price. Instead of $\bar{p}$ we might have just as easily chosen any other price level. Accordingly, what has been said about the point $(\bar{p}, \bar{x})$ on the demand curve can be said for any other point on this curve. Generally, the total demand curve is constructed by adding up the quantity values of all of the individual demand curves for any price value. Since the quantity values to be added up are read from the abscissa, this procedure is called the generation of the total demand curve by "horizontal aggregation" of the individual demand curves.

Please note that even though the two individual demand curves in the figure are linear, the aggregate demand curve has a kink. This is due to the fact that the prohibitive price, $b$, of the second consumer is higher than the prohibitive price of the first consumer, $a$. So for prices between these two prohibitive prices, consumer 1 does not contribute any quantity demanded in the process of horizontal aggregation of the two individual demand curves. For prices between $a$ and $b$, the aggregate demand curve is identical to the individual demand curve of the second consumer. It is only if the prices used for the horizontal aggregation are lower than $a$ that positive quantities demanded by the first consumer are added to the demanded quantities of the second consumer.

Analogously, the market supply curve is derived from the individual supply curves of all of the firms operating in the market under consideration, which is also achieved by horizontal aggregation. Because of the analogy, we present only the graphical illustration, leaving comments brief.

We illustrate the supply curves of the two firms, $s_1(p)$ and $s_2(p)$. Then, we arrive at the total (aggregate) supply curve, $S(p)$, by horizontal aggregation. For each given price, the individual equilibrium supply quantities of the two firms are added up; their sum is the total quantity supplied. The combination of the predetermined

[43] In more formal terms we may write $i \in \{1, 2\}$.

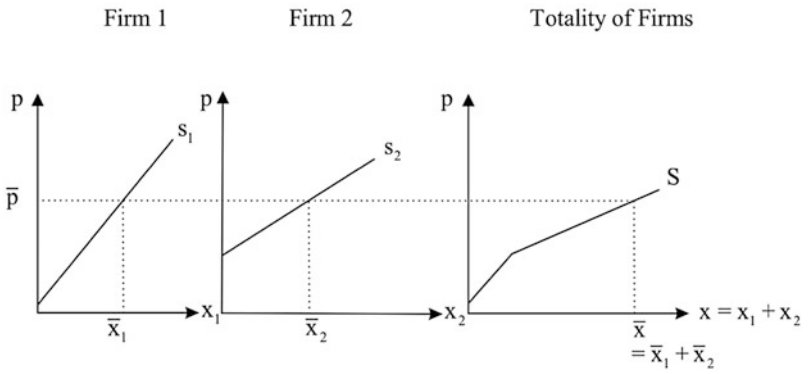


**Fig. 6.8** Deriving the market supply curve by *horizontal* aggregation of individual supply curves

price and the corresponding total quantity supplied is one point on the total supply curve. Combining all of the points generated by this procedure ("drawing a line through those points in the graph") constitutes the total supply curve, $S(p)$ (Fig. 6.8).

Now we have designed all the pieces to be able to put the puzzle together. To illustrate how demand meets supply in the market, we draw the two previously derived curves, $S(p)$ and $D(p)$, into one diagram. This diagram is probably the most often drawn graphic in economics classes and textbooks.[44] Some have said that it constitutes the *totem* of the "tribe of the econ", the members of the economics profession (Fig. 6.9).

It is immediately obvious that the point where the two curves intersect requires special attention. Given the price is at $p^*$, the total quantity demanded and the total quantity supplied are identical to each other at $x^*$. In this situation the market is said to be in *equilibrium*. $p^*$ is the *equilibrium price* and $x^*$ is the *equilibrium quantity*. It should be recalled that the individual demand curves underlying the market demand curve and the individual supply curves underlying the market supply curve have been derived from utility maximization of each consumer and from profit maximization of each firm, respectively. Resultantly, it holds that in the market equilibrium each consumer maximizes utility and each firm maximizes profit, simultaneously. In addition, these equilibria of the individual agents are compatible with each other.[45]

To further highlight the idea of market equilibrium, let us briefly consider a situation of disequilibrium. To do so, imagine a price above the equilibrium price at, say, $\bar{p}$. Given this price, utility maximization of the consumers would lead to a market

[44] It is also one of the "famous figures and diagrams" beautifully explained in Blaug and Lloyd (2010). The authors did the right thing to put the supply and demand diagram referred to above right at the beginning of their gallery of fundamental economic illustrations.

[45] In this situation every agent is able to realize his/her plan.

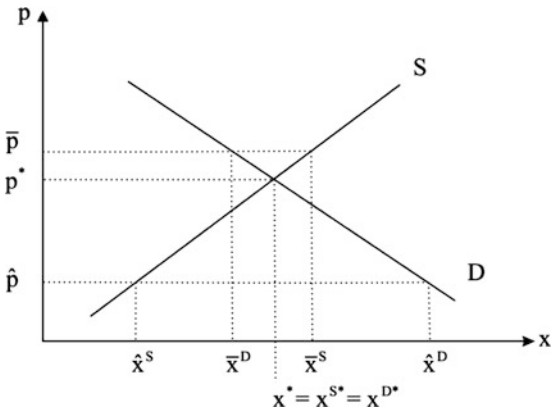


**Fig. 6.9** Market equilibrium

demand of $\bar{x}^D$ units of the good under consideration. On the other hand, the production plans of all the firms would lead to a total quantity supply of $\bar{x}^S$. However, the plans of the consumers to maximize their utilities and the plans of the firms to maximize their profits are not compatible with each other, given the price is at $\bar{p}$. $\bar{x}^S$ is bigger than $\bar{x}^D$, such that firms are not able to sell $\bar{x}^S - \bar{x}^D$ units of the product at the price of $\bar{p}$. This situation in which the producers would be sitting on parts of what they would have produced is called *excess supply*. It is intuitive that a situation of excess supply might generate a downward pressure on the price, i.e., in the direction of the equilibrium price. This would stimulate the incentive to consume and simultaneously attenuate the incentive to produce, thereby narrowing excess supply. Under ideal conditions, the process would go on until equilibrium is achieved.[46]

The analogous argument holds if we observe a price below the equilibrium price, at, say, $\hat{p}$. Given $\hat{p}$, the quantity demanded in this market would be higher than the quantity supplied, resulting in a situation of *excess demand*, $\hat{x}^D - \hat{x}^S$. Here, the utility maximization and profit maximization decisions of the agents would not be compatible with each other in that the consumers would not be able to buy the quantities they are planning to, at this price. It is "plausible" that the price would increase in this situation and thereby would move into the direction of the equilibrium price. Consequently, quantities demanded go down and quantities supplied go up, narrowing excess demand (ideally) until it vanishes and thereby equilibrium is achieved.[47]

[46] Among the ideal conditions mentioned above is the requirement that all agents are fully aware of ruling prices. Moreover, it is a prerequisite that no governmental intervention prevents prices from reacting to divergences between the quantity supplied and the quantity demanded.

[47] The economic models presented above have been confined to static and comparative static equilibrium analysis. Therefore, strictly speaking, the speculations about how prices would move starting from a situation of disequilibrium violate the limits of our analysis. Therefore we intentionally use the soft term "plausible" in the passage to which this footnote refers.

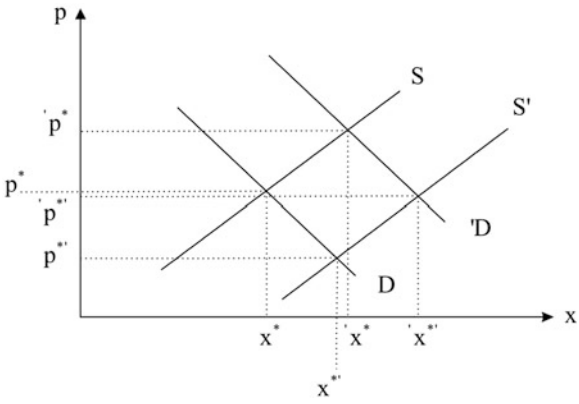


**Fig. 6.10** Comparative static analysis of the market equilibrium

We have argued that it is plausible to expect the price to move towards the equilibrium price if the price is too high or too low in the initial situation, i.e., at $\bar{p}$ or $\hat{p}$ in the illustration presented above. In case this is true, the equilibrium under consideration is said to be *stable*. However, the stability of market equilibrium is not a natural law. There may be cases where actual prices are driven further away from the market equilibrium price if the starting price is not identical to the equilibrium price; in such cases, the market equilibrium is said to be *unstable*.[48]

In the main part of the analysis given above, we have commented on the market equilibrium in terms of a *static analysis* (enriched with just a glimpse at economic dynamics). Supply and demand curves did not change in the process of the analysis. However, in respectively dealing with the individual consumer and the individual firm, we previously briefly discussed *comparative static analysis*. In the case of the consumer we used an increase in income as an example to show that the demand curve might shift. In the case of the firm we used changes in the prices of the inputs to the production process, as well as advances in production technology, as two examples to show that the supply curve might shift. Of course, these insights can be transferred from the level of the individual decision maker to the level of the market. Essentially, this suggests that if the income of all consumers changes, the market demand curve shifts. If input prices change or production technology improves, for all of the firms, the market supply curve shifts. We illustrate this with the supply curve moving from $S(p)$ to $S'(p)$ and the demand curve moving from $D(p)$ to $'D(p)$ (Fig. 6.10).

You can see that the downward shift of the supply curve due to technical progress (taking the demand curve as it is, at $D(p)$) induces a decrease in the equilibrium price from $p^*$ to $p^{*\prime}$ and an increase in the equilibrium quantity from $x^*$ to $x^{*\prime}$. An outward shift of the demand curve from $D(p)$ to $'D(p)$ due to an increase in income (leaving the supply curve at $S(x)$) induces an increase in the equilibrium price from $p^*$ to $'p^*$and an increase in the equilibrium quantity from $x^*$ to $'x^*$. If we take the changes in demand and supply to occur simultaneously, i.e., we

[48] The question of the stability of market equilibria is a key topic of *dynamic* economic analysis.

assume that the supply curve shifts from $S(x)$ to $S'(x)$ and the demand curve shifts from $D(x)$ to $'D(x)$ "at the same time",[49] then, equilibrium price changes from $p^*$ to $'p^{*'}$ and equilibrium quantity changes from $x^*$ to $'x^{*'}$. Please note that it cannot be said in general terms whether the new equilibrium price is above or below the old one. This is due to the fact that the two changes introduced have countervailing effects. The improvement in production technology creates a tendency to bring the price down. On the other hand, the increase in income creates a tendency to drive the price up. The net effect depends upon which of these countervailing powers is stronger. Analogous reasoning holds for the effect of simultaneous changes in the demand and the supply curve on the equilibrium quantity.

## 6.5 Basic Issues in Welfare Economics

### 6.5.1 The Concept of Social Optimality

In the previous sections we dealt with the principle actors on the stage of microeconomics, consumers and firms, and with the script according to which these actors play together, the market mechanism.

This was all positive analysis: we described how individual decisions and their coordination are stylized in microeconomics. We did not ask whether the results of these decision and coordination processes are good or bad for society as a whole. In the present section we are going to consider this normative issue. Doing so we take up what we briefly touched on in the paragraph on "microeconomics as a theory of the evaluation of resource allocation" in Sect. 6.1, above. There, we promised that the idea of operationalizing the general concept of social welfare would be elaborated on in subsection 6.5.1. So now is the time to deliver!

The measuring rod with which it may be assessed what is good or bad for society as a whole is called a *social welfare criterion*. If social welfare goes up according to such a criterion, this is taken to be good for society, and vice versa.

One very popular social welfare criterion has been around for a particularly long time: it was introduced by the Italian sociologist Vilfredo Pareto (1848–1923), and to honour his contribution it is called the *Pareto criterion* in the microeconomics literature. It says that an allocation A is better than an allocation B for society as a whole if in A at least one member of society is better off than in B, and no other member of the society is worse off in A than in B.[50]

[49] We put this expression in quotes because, strictly speaking, there is no time in the model that we use here. To allow for that, we need a dynamic model but the one under consideration is static. So of any two activities it is not possible that one happens "first" and the other happens "second", nor can they happen "at the same time". However, it supports intuitive grasp on what is going on in the model if we talk about it using terms from daily "colloquial" language, even if they are not strictly appropriate.

[50] Whether the particular individual is better off in A or B is decided by this very individual. This follows from the principle of consumer sovereignty, referred to above.

Take a group of people sitting in a (class)room together. The group in this little example substitutes for "society" in the general exposition given above. The window in the room is closed, which we notate as allocation $A$. Now consider changing to allocation $B$, which is identical to $A$, except for the window being open. If, in the first situation, some members of the group (at least one of them) want the window to be opened and no one objects, it is a Pareto improvement to open the window (to go from $A$ to $B$).

Developing the idea a little further, this criterion does not only allow us to say which one of two alternative allocations is better for society as a whole. In addition, it allows us to identify allocations which cannot be improved from the point of view of society as a whole. An allocation which cannot be improved in this sense is called *Pareto optimal*. In a Pareto optimal allocation it is not possible to introduce any change by which the utility of at least one member of society is increased without making any other member of society worse off.

According to the experience of the authors, the Pareto criterion is plausible and acceptable to most people. Of course, students with whom we discussed the acceptability of the Pareto criterion are not representative of the world population (and the students of other economists who had the same experience aren't either). Still, the acceptance that we have observed is not very surprising. After all, there are similarities between deciding whether a certain reallocation of resources is an improvement for society as a whole according to the Pareto criterion and to deciding this issue by unanimous vote: if, as a result of a certain reallocation, some people would enjoy higher utility levels and no other people would suffer from a deterioration in their utilities, then you might expect that the people who would benefit from the change would vote in favour of the reallocation and others would not vote against it, say abstain.[51] However, it must be conceded that the Pareto criterion is not "value free". After all, the criterion implies that the welfare of society is nothing more than the welfares of the individual members of society. This value judgement is called *normative individualism*, in economics. Even though this is quite a common perception, you cannot prove anybody wrong who believes otherwise.

The fact that the Pareto criterion is widely (even though not "mandatorily") accepted is one of its advantages. On the other hand, the criterion is not very powerful in its ability to help societies make decisions between alternative allocations. This is so because an awfully high number of allocations meet the criterion of Pareto optimality.[52] Without referring to complicated proofs, this is intuitively clear when you recall the analogy between the Pareto criterion and unanimous voting. There are very many situations from which it is impossible to deviate by a unanimous decision of the members of the society under consideration. All of these situations are Pareto optimal.

[51] The assumption underlying this expectation is that each individual votes according to his/her own self interest. This implies that the people who are not affected by the reallocation do not envy the ones who benefit.

[52] Strictly speaking there are an infinite number of allocations. See intermediate microeconomics textbooks for deeper analysis. Examples are Eaton et al. (2011), Varian (2010).

To illustrate the point from our earlier example, imagine a situation where the group enters the classroom and the window is closed, i.e., the starting situation is *A*. It might happen that at least one member of the group wants the window to be opened but there is at least one other member of the group objecting. Then, the starting situation with the closed window is Pareto optimal. Obviously, it is not possible to change the situation and make at least one member of the group better off without making another member of the group worse off.

Alternatively, imagine the group entering the room and finding the window open (starting at situation *B*). It is not contradictory to the story told above that at least one member of the group wants to have the window closed but at least one other member objects. If this is the starting situation, it is Pareto optimal because you cannot go from *B* to *A* making at least one member of the group better off without making at least another member worse off. Our example therefore illustrates a case where the choice is between two allocations that are both Pareto optimal. Obviously in such a situation the Pareto criterion is not a very helpful social decision rule.

There have been attempts to broaden the set of situations which can be ranked according to their social desirability by modifying the Pareto criterion. The essence of these approaches is that we take a look at reallocations from which some members of society gain and others loose. There, the winners would vote in favour of the reallocations and the others would vote against it. So the change under consideration could not be assessed using the Pareto criterion. To overcome this problem you would look at *how much* the people who benefit from the change would win and *how much* the people who suffer would loose. According to the *Kaldor-Hicks criterion,*[53] the change would be a social improvement if what the people benefiting won would be enough to compensate the losers and still leave a positive net improvement for the ones who would benefit from the change in the first place. Please note that, according to this criterion, it is not a prerequisite that the compensation actually takes place. All that is required for the reallocation to qualify as a social improvement is that compensation is possible. (If the compensation does, indeed, take place we are back to the original Pareto criterion.) Therefore the Kaldor-Hicks criterion is often termed a criterion of *potential* Pareto improvement.

A related approach to assessing the desirability of alternative allocations from the point of view of society is the *social welfare function*. This one comes in different variants.

The most simple type is the *utilitarian social welfare function*. Here, the welfare of society is defined as the sum of the utilities of the members of society. All of these individual utilities contribute to social welfare with the same weight. A generalization of this simple concept of aggregation is the *Bergson-Samuelson welfare function*.[54] Here, social welfare is the weighted sum of individual utilities.

[53] Nicholas Kaldor (1908–1986), British economist; John R. Hicks (1904–1989), British economist, Economics Nobel Prize 1972.

[54] Abram Bergson (1914–2003), US-American economist; Paul E. Samuelson (1915–2009), US-American economist, Economics Nobel Prize 1970.

This creates some leeway for societal welfare judgements. For example, the utilities of the "weaker" members of society might enter into the aggregation with a higher weight. Obviously, the Bergson-Samuelson welfare function contains the utilitarian welfare function as a special case – the one where all the weights are equal to 1. An obvious problem is that defining the weights is difficult and opens up considerable room for all kinds of discrimination. A third prominent type of a social welfare function is the one designed by the US-American philosopher John Rawls (1921–2002) in his monumental work "A Theory of Justice" (1971). Here, social welfare is determined by the utilities of those members of society who are worst off, only. As you might already have suspected, this special view of justice has been highly controversial in the literature.

Comparing the Pareto criterion to the social welfare function approach, we might note that it is an advantage of the latter that all alternative allocations can, in principle, be ranked according to their social desirability. Conceptually it is thereby possible to identify a unique socially optimal allocation (provided it exists). This implies the possibility of arriving at very definite suggestions as to where a society should move. As we have argued above, this is impossible using the Pareto criterion. However, as very often in life, "there is no such thing as a free lunch": using the social welfare function approach we have to pay a price for enjoying the described advantage. We must assume that we can cardinally measure the utilities of all the individual members of a society in the same dimension. Consider the utilitarian welfare function as an example. In order to add up all the individual utilities to arrive at a number for social welfare, it must be possible to measure utilities just as you measure distances in yards or metres. The problem is (and this is a considerable problem, indeed) that such a quantitative measure of the utilities of different people does not exist.

You might suppose that, in light of this impossibility, microeconomists would have given up on the idea of a social welfare function. Far from it! Microeconomists do not give up that easily. Instead, they have been looking for a measuring rod which might be substituted for the unattainable goal of measuring utility cardinally. What microeconomists have come up with is *willingness to pay*. Willingness to pay certainly meets the requirement of being quantifiable cardinally in one and the same dimension for different individuals. This dimension is money. Also, there is certainly a strong positive correspondence between the utility an individual expects to derive from a certain reallocation (e.g., by receiving a certain commodity) and his/her willingness to pay for this reallocation. For most people the willingness to pay for a certain good increases with the level of utility they expect to attain by consuming the good. However, even though they are closely related to each other, utility and willingness to pay are not the same. This is so because the willingness to pay not only depends on the preferences of an individual but also on the *ability to pay*, i.e., the size of the budget the individual has at his/her disposal. A wealthy person can articulate his/her utility much better in terms of willingness to pay than a poor person. So if aggregate willingness to pay is used as a proxy for social welfare, this basically applies the Bergson-Samuelson type of welfare function in a specification where the weight with which the utility of an individual enters the social welfare function increases with the income of this individual. This type of a social

welfare function is suggestive only in societies where there is a consensus that the income distribution is (well, more or less) just.[55]

Imagine that in a society there is a distribution of income that is understood to be just in a pragmatic sense. This does not mean that everybody in this society is happy with the distribution. It would be compatible with our understanding of a distribution "being just in a pragmatic sense" if most people thought that the effective distribution was somewhat unfair, in the sense that they themselves should have received a little more and the others a little less. However, the effective distribution is the result of market allocation and subsequent redistributional government policy. The redistributive activities of the government like progressive income taxation, the social welfare and education systems, and others have been decided upon by a democratic process. In light of this democratic legitimization it can be said that the society agreed upon accepting a certain distribution, even if some members of the society might have done so with clenched teeth. In such a situation it might be reasonable to use aggregate willingness to pay as a proxy to social welfare – unless you can come up with something better. This is then a powerful tool for evaluating different allocations according to their social desirability.

Armed with what we have just said, we are able to assess the results of individual decision making and its coordination from the point of view of society as a whole. This is what we do in the following subsection, applying the concept of social optimality which has just been operationalized to the performance of consumers, firms and the market mechanism, which were explained in the previous sections.

### 6.5.2 The Social Optimality of an Ideal Market Economy

#### 6.5.2.1 Socially Optimal Output

In Sect. 6.4, above, we conducted a strictly *positive* analysis characterizing the equilibrium quantity, $x^*$, of an arbitrarily chosen good, X, which is provided in a perfectly competitive market. In subsection 6.5.1, above, we conducted a strictly *normative* analysis, pondering the concepts of social welfare and social optimality. Now we put the pieces together, and assess the social welfare properties of the perfectly competitive output quantity. Applying the concept of social optimality to the problem of what quantity of a certain good should be produced leads to the idea of the quantity that maximizes social welfare. The social welfare of production is the benefit this production yields for the consumers minus the opportunity costs of the resources used to produce the quantity of that good. Using this operationalization of the idea of social welfare, we heavily rely on the concept of *willingness to pay*, explained in the previous section.

[55] The understanding of what fairness (justice) is varies over time and across societies with different cultural backgrounds.

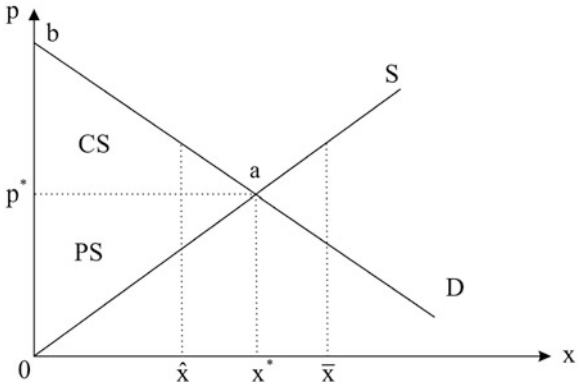


**Fig. 6.11** Socially optimal and perfectly competitive level of production

It is plausible that these two elements, consumers' benefits and opportunity costs, are exactly the driving forces identified above as constituting the equilibrium of a perfectly competitive market.

The benefits to consumers are "caught" in the demand function. This is so because a monetary measure for these benefits is the willingness to pay of consumers. As has been explained in Sect. 6.2, marginal willingness to pay can be read off the inverse demand curve. So given that the ordinate value of each point on the inverse demand curve shows the willingness to pay for the last ("marginal") unit consumed, then the area under the demand curve illustrates total willingness to pay, i.e., the benefits from consumption.

Analogously, it can be argued that the inverse supply curve illustrates the opportunity costs of the resources used up by the production. This is so because the inverse supply curve is generated from the marginal cost curves of the individual firms. So at each point the ordinate value of the inverse market supply curve represents the production cost of the last ("marginal") unit produced. Accordingly, the area under the inverse supply curve is the total cost of the resources used for production (ignoring fixed cost).

Following the goal of social welfare maximisation, production of any unit under consideration is warranted if marginal willingness to pay is at least as high as marginal production cost. In the terminology of the graphical illustration used in this text, such as in Fig. 6.11, this means that the production of a certain unit of $X$ is socially warranted if the ordinate value of the inverse demand curve is at least as high as the ordinate value of the inverse supply curve. This criterion for the social desirability of a certain unit of $X$ is met for all the units between 0 and $x^*$, the level of output at which demand and supply curves intersect. This is exactly the perfectly competitive equilibrium. Take any other unit of production (say, $\bar{x}$, where $\bar{x} > x^*$), as indicated in the figure. For this quantity the benefit to the consumers is lower than the value of the resources used to produce this unit. In $\bar{x}$, the inverse demand curve is below the inverse supply curve. Therefore, this unit should not be produced

according to the criterion of social welfare maximization. Indeed, it is not produced in the perfectly competitive equilibrium.

On the other hand, if production would be at a level of $\hat{x}$, smaller than $x^*$, then the production of the last unit would be socially warranted. However, the possibility of increasing social welfare would not be exhausted. This is because between $\hat{x}$ and $x^*$ there is plenty of room to produce units of the product for which the marginal willingness to pay is above marginal production cost. In the figure, the level of social welfare generated by the optimal and equilibrium production $x^*$ is illustrated by the area between the inverse supply and the inverse demand curves, within the limits of 0 and $x^*$. This is the area $0ab$.

In the figure, you can also read how the social welfare generated is shared between consumers and producers. The total benefit enjoyed by the consumers from the socially optimal and perfectly competitive quantity $x^*$ is measured by the area under the inverse demand curve within the limits of 0 and $x^*$, the area $0x^*ab$. However, to get hold of $x^*$ in the market, consumers must pay the equilibrium price, $p^*$ for each unit of the product. As a result, they spend a sum equal to $p^*x^*$, illustrated by the area $0x^*ap^*$ in the figure. Of course, this reduces their overall benefit. The net benefit they receive is the willingness to pay minus the sum actually paid, i.e., an amount of money represented by the area under the inverse demand curve and above the horizontal price-line within the limits of 0 and $x^*$, $p^*ab$. This net benefit the consumers gain from consumption is called *consumer surplus* ("*CS*" in the figure).

A similar argument can be made for firms. The benefit of producing is the revenue they attain in the market. What they get is identical to what the consumers pay (this is particularly so in the present model, since there is no one collecting taxes). Ergo, in the illustration, the benefit to the producers is $p^*x^*$, the area $0x^*ap^*$. On the other hand, the producers must pay for the scarce resources they use in the production process. This amount of money is illustrated by the area under the inverse supply curve between 0 and $x^*$, i.e., by $0x^*a$. The net benefit to the producers is graphically represented by the area between the horizontal price line and the inverse supply curve between 0 and $x^*$, i.e., by $0ap^*$. This net benefit is called *producer* surplus ("*PS*" in the illustration).

Consumer surplus and producer surplus add up to the social welfare generated by production. You might as well call the latter "social surplus". Consequently you might also say that in a perfectly competitive equilibrium the quantity produced is socially optimal, since for this quantity the sum of consumer and producer surplus is maximized.

The result that the equilibrium of a perfectly competitive economy is socially optimal (as far as shown here: in terms of output levels) is of utmost importance for economic theory. The result is therefore terminologically knighted with the expression "the first fundamental theorem of welfare economics".[56]

[56] Obviously, there would not be a "first" theorem if there were no "second" one (and possibly others). However, to discuss this would lead us astray from the convenient pathways of our

The result is somewhat surprising. After all, there is no economic agent in the model striving for social welfare maximization. On the contrary, each of the agents under consideration is exclusively interested in his/her own welfare. This individual welfare takes the form of utility in the case of a consumer and of profit in the case of a firm. Instead of being attained by the activities of a superior actor, e.g., a "social planner" the result is attained by the "invisible hand" of the perfect market guiding the individual actors.[57]

Moreover, the allocative mechanism of the perfectly competitive market does not need an outrageous amount of information to do the *Herculian* job of producing socially optimal equilibria. All that is required is that each individual actor knows the market prices of the goods under consideration (commodities, services, and productive inputs) and their qualities. Additionally, actors who are consumers need to know their preferences and budgets. Actors that are firms need to know production technology. It is worth noting that no individual consumer is required to know the preferences of any other consumer or the technology of any firm. No individual firm is required to be informed about the technology of any other firm or about the preferences of any consumer.

If the task of producing socially optimal output would have to be fulfilled by a central planning agency, this agency would have to look behind the foreheads of all the consumers to learn about their preferences and behind the walls of the firms to learn about technology. Certainly, the consumers and the firms would stage some resistance against this.

#### 6.5.2.2 Cost-Effective Inter-firm Allocation

What we said above on the social optimality of equilibrium output in a perfectly competitive economy is what you generally see in introductory economic textbooks. Fine. However, the authors of the present text just love to coddle their readers. So here is an additional and somewhat more subtle argument contributing to a deeper understanding of the welfare properties of perfectly competitive equilibria. It will turn out to be useful when we consider the economics of environmental policy in Sect. 7.2, below.

We have argued above that a prerequisite for a firm to achieve its goal of maximizing profits is that it minimizes cost. This follows from the definition of profit as the difference between revenue and cost. No matter how much a firm decides to produce, without cost minimization there is no profit maximization. If a given quantity of a good is produced at minimum cost, production is said to be *cost-effective*.

This kind of reasoning can be transferred from the individual firm to society as a whole. As we have discussed above, a standard assumption in microeconomics is

introductory exposition. See, e.g., Estrin et al. (2008), pp. 468–476, Perloff (2007), pp. 318–321, Varian (2010), pp. 601–606.

[57] The *invisible hand* is a "crossover word" going right back to the cradle of modern economics, the work of Adam Smith (1776).

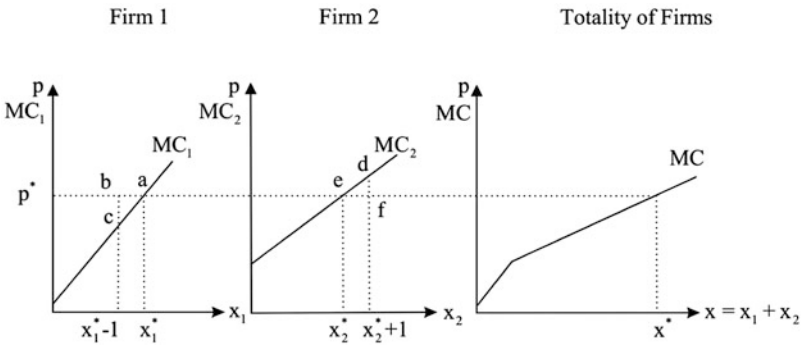


**Fig. 6.12** Cost-effective inter-firm allocation

that society strives for welfare maximization. Welfare is defined as the difference between aggregate consumer benefit and aggregate cost.

A prerequisite for social welfare maximization is that the costs of production (represented in the inverse supply function in marginal terms) are minimized.

On the level of the individual firm this is no problem since each firm is interested in cost minimization, as has been explained above. However, the market supply curve is generated from individual supply curves by a special procedure: horizontal aggregation.[58] The question is whether this procedure satisfies the requirement of cost minimization. Wouldn't it be possible that the market mechanism makes the individual firms contribute to total supply in a manner that violates the cost minimization principle? Wouldn't it be possible to "assign" shares of total output to the individual firms where those shares are different from the distribution as generated by the market mechanism and in doing so arrive at the production of total output at lower cost?

We use Fig. 6.12 to illustrate.

Assume the equilibrium market price is at $p^*$. Then, equilibrium total supply is at $x^*$. To this total quantity, firm 1 contributes $x_1^*$ and firm 2 contributes $x_2^*$ units, where $x_1^*+x_2^* = x^*$ holds. However, $(x_1^*,x_2^*)$ is just one of indefinitely many inter-firm allocations of the total supply of $x^*$. $(x_1^* - 1,\ x_2^* + 1)$ is an alternative allocation which sums up to exactly the same level of aggregate output, $x^*$. A question is whether the allocation $(x_1^*,x_2^*)$ brought about in the perfectly competitive market equilibrium is the cost minimizing allocation.

Yes, it is!

To see why, we make a comparison between the cost of producing $x^*$ with firm 1 contributing $x_1^*$ and firm 2 contributing $x_2^*$ with the cost of producing $x^*$ with firm 1 contributing $x_1^* - 1$ and firm 2 contributing $x_2^* + 1$.

You can do that by staging a little thought experiment: imagine the allocation is $(x_1^*,x_2^*)$ and you figure out that the total cost of producing $x^*$ can be reduced by

[58] That's exactly what we did in Sect. 6.4, above. See Fig. 6.8.

changing the allocation to $(x_1^* - 1,\ x_2^* + 1)$. If total cost decreases as a consequence of this reallocation, then – gotcha! – the market equilibrium $(x_1^*,x_2^*)$ cannot be a cost minimizing allocation.

So what are the consequences of going from $(x_1^*,x_2^*)$ to $(x_1^* - 1,\ x_2^* + 1)$ in terms of total cost?

The cost reduction at firm 1 due to the decrease of this firm's production from $x_1^*$ to $x_1^* - 1$ is illustrated by the area under the marginal cost curve of firm 1 between $x_1^* - 1$ and $x_1^*$, i.e., by the area $(x_1^* - 1)x_1^*ac$.

To compensate for the unit less produced by firm 1, in terms of the level of total output, firm 2 has to produce one unit more. The additional cost is illustrated in the figure by the area under the marginal cost curve of firm 2 within the limits of $x_2^*$ and $x_2^* + 1$, i.e., by the area $x_2^*(x_2^* + 1)de$.

Obviously, the amount by which the costs of firm 2 increase is higher than the amount by which the costs of firm 1 decrease. Thereby, the sum of the costs the two firms have to bear, producing an aggregate quantity of $x^*$, increases by an amount graphically illustrated as the sum of the two triangles *abc* and *def*.

Deviating from the market equilibrium inter-firm allocation $(x_1^*, x_2^*)$ to the alternative allocation $(x_1^* - 1,\ x_2^* + 1)$ would increase the cost burden society has to bear for the benefit of producing $x^*$. Therefore, the reallocation that we have considered would fail the cost-effectiveness test. Of course, we have analyzed only one of indefinitely many possible deviations from the market equilibrium inter-firm allocation. However, the principle that we used in discussing this example is perfectly general. Starting from the market equilibrium allocation you can take any reallocation. You will find that the result is always the same. Deviating from how the perfectly competitive market "assigns" the shares to which individual firms contribute to total equilibrium supply increases the burden society has to bear for the sake of production. So the inter-firm allocation of total output, as "arranged" in the perfectly competitive market equilibrium, is cost-effective. Thereby, an important prerequisite for social welfare maximization is met in equilibrium.

The deeper reason for the result graphically illustrated above, is that given that marginal cost curves are increasing, it is a property of the cost minimizing allocation that the marginal costs of different firms are equal to each other. The "trick" of the perfectly competitive market is that in its equilibrium this property of cost minimization is met. The mechanism that guarantees this is profit maximization, while also influential is the fact that an individual firm has no command over the level of the market price under perfect competition (and that this market price is identical for all of the firms producing the same good).

As we have argued in the preceding section (on the firm), a firm (let's call it "i") maximizes profit by choosing to produce the level of output for which its marginal cost is equal to the market price, $p^* = MC_i$.[59] This is true for each of the firms supplying in the perfectly competitive market. Since the market price is the same

[59] In our 2-firms-example, "i" might take the values of 1 or 2. Some highbrows might write $i \in \{1, 2\}$.

for all of these firms, marginal costs are identical across firms, in the competitive equilibrium. So $MC_i = MC_j$ holds for any pair of firms.[60] The equality of the marginal costs of the firms is a requirement for cost minimization and is simultaneously a feature of the perfectly competitive equilibrium.

**Calculus Club: Session 3**

The problem of cost-effective inter-firm allocation of a predetermined quantity of aggregate production, $\bar{x}$, has been dealt with above using verbal and graphical analyses. It can be laid out in formal terms as follows:

The objective is to minimize total production cost, $C(x)$, which is defined as the sum of firm-specific production costs, $C_1(x_1)$, $C_2(x_2)$. We do this for two firms here without any loss of generality. The formal expression of this cost minimization problem is

$$C = C_1(x_1) + C_2(x_2) = \min!$$

Production cost minimization is subject to the requirement that output produced by the two firms, $x_1$ and $x_2$, adds up to an aggregate output of $\bar{x}$. So the minimization has to be done under the constraint of

$$x_1 + x_2 = \bar{x}.$$

Consequently, the Lagrange function is

$$L = C + \lambda(\bar{x} - x_1 - x_2) = \min!$$

Writing the first order derivative of the total cost function, the marginal cost function, as MC, the first order conditions are

$$\begin{aligned} \partial L/\partial x_1 &= MC_1 - \lambda = 0 \\ \partial L/\partial x_2 &= MC_2 - \lambda = 0 \\ \partial L/\partial \lambda &= \bar{x} - x_1 - x_2 = 0 \end{aligned}$$

$$\rightarrow MC_1 = MC_2.$$

So the necessary condition for the cost-effective inter-firm allocation we are looking for is that marginal production costs are equal across firms.

Since the second cross dervivatives equal zero, the second order conditions are

[60] If "i" is one of the two firms in our example, "j" is the other one.

$$\partial^2 L/\partial x_1^2 = \partial MC_1/\partial x_1 > 0$$
$$\partial^2 L/\partial x_2^2 = \partial MC_2/\partial x_2 > 0\cdot$$

Accordingly, the extreme value characterized by the first order condition is indeed a minimum if it is located in the increasing parts of the two marginal production cost curves. Since, for simplicity, we assumed the marginal production cost curves to be monotonously increasing, above, this second order condition is always met.

#### 6.5.2.3 "Benefit-Effective" Inter-consumer Allocation

Above, we have compared the cost-effective inter-firm allocation of a given level of aggregate production to the perfectly competitive inter-firm allocation. The result of the analysis is that they are identical.

Obviously, an analogous question can be posed regarding the manner in which a predetermined quantity of aggregate production is allocated to the consumers. What would be the "right" way to share total production among the consumers of the economy under consideration and how does this "right" scheme relate to the perfectly competitive scheme?

It may not be immediately clear what is meant through the nobly termed criterion for the allocation among consumers of "right". Every reader might have his/her attitudes and thoughts about this question. However, in the present context we are not dealing with personal attitudes and thoughts but with microeconomic theory.[61] Indeed, microeconomics offers an answer to this difficult question. All we have to do in order to find out is to be consistent with the general concept of social welfare maximization. Above, social welfare has been defined as the difference between aggregate willingness to pay and opportunity cost. We used this definition above to find out about the "right" inter-firm allocation of a certain aggregate production. There, the focus was on the cost side of the social welfare definition. In the present context, where we deal with the "right" allocation of a certain aggregate production among the consumers, we proceed analogously by focusing on the side of aggregate willingness to pay. Consequently, the "right" allocation among the consumers is the one that maximizes aggregate willingness to pay for the given production level because willingness to pay is a measure of the benefit from consumption.[62] With reference to the terminology used with regard to firms, terming the "right" inter-firm allocation as cost-effective, we call the inter-consumer allocation generating maximal aggregate benefit as a "benefit-effective" allocation.

[61] Nevertheless, every reader is cordially invited to compare his/her own attitudes and thoughts with microeconomic wisdom.

[62] However, we remind our cherished readers of our discussion of the caveats of this concept, as presented above.

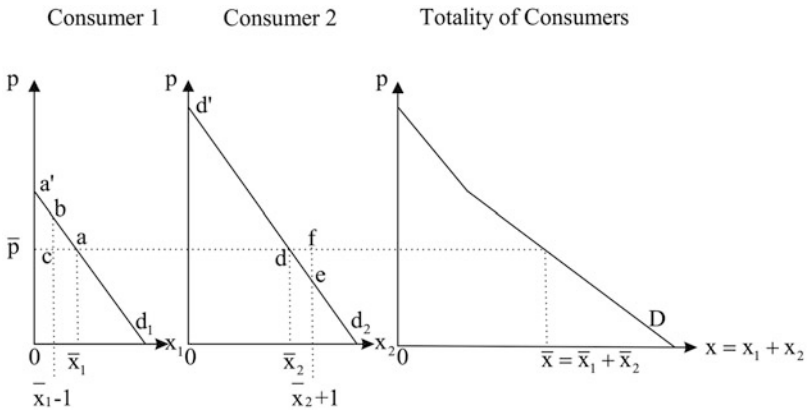


**Fig. 6.13** "Benefit-effective" inter-consumer allocation

Also in perfect analogy to what has been said for the firms, above, the benefit-effective allocation is characterized by the requirement that marginal values for the agents under consideration have to be equal to each other. With regard to firms, these marginal values have been shown to be marginal costs. With regard to the consumers, it's each consumer's marginal willingness to pay for the good under consideration that has to be equal in the benefit-effective situation. To see this, take a look at Fig. 6.13.

Here an inter-consumer allocation is shown where the aggregate production quantity, $\bar{x}$, is shared among the two consumers in that consumer 1 receives a quantity of $\bar{x}_1$ and consumer 2 receives $\bar{x}_2$.[63] The benefit a consumer derives from consumption is measured by the area under the inverse demand curve in the relevant range. In the situation described above, the benefit to consumer 1 is illustrated by the area under the inverse demand curve $d_1$ within the limits $0\bar{x}_1$, i.e., by the area $0\bar{x}_1 aa'$. Analogously, the benefit of consumer 2 is illustrated by the area $0\bar{x}_2 dd'$.[64] It can be shown that this allocation, characterized by the marginal willingness to pay of the two consumers being identical to each other, generates the maximum benefit that can be "squeezed out" of total production, $\bar{x}$. To demonstrate this, follow the analytical line of reasoning explained above. You can show that aggregate benefit for the two consumers decreases if you deviate from the aforementioned allocation. As an example for such a reallocation let consumer 1 have one unit less and consumer 2 one unit more. Then, track down the consequences of this reallocation in terms of aggregate benefit. It is easy to see that what consumer 2 gains from having one unit

[63] Please remember that the total demand curve shown in Fig. 6.13 is the horizontal aggregation of the two individual demand curves shown in this figure. The procedure has been explained in Sect. 6.4, above, using Fig. 6.7.

[64] Please avoid a notational trap here: do not confuse the points on the demand curve, denoted "$d$", "$d'$", with the demand curve labelled $d_2(x_2)$

more is less than what consumer 1 loses from having one unit less, in terms of benefits. The total decrease in the aggregate benefit due to the reallocation under consideration is illustrated by the sum of the areas *abc* and *def* in Fig. 6.13.[65]

By taking any other reallocation starting from $\bar{x}_1$, $\bar{x}_2$, the result will be, in principle, the same: aggregate benefit goes down. So the "golden rule" for allocating a given amount of production among the consumers is to do it in a way that ends up with the consumers having identical marginal willingness to pay.

Obviously, this requirement is met by the perfectly competitive allocation. There, the equilibrium is characterized by the fact that each consumer buys a quantity at which his/her marginal willingness to pay equals the market price, and this price is the same for all the consumers. Accordingly, marginal willingness to pay is the same across consumers in equilibrium. From $\bar{p} = MWP_1$ and $\bar{p} = MWP_2$, $MWP_1 = MWP_2$ follows, where MWP stands for the marginal willingness to pay.

### 6.5.3 Market Failure

In the previous subsection we praised the social welfare properties of a perfectly competitive market. This kind of market is a useful analytical and didactical tool. However, it does not adequately represent the world in which we live. Instead, there are many systematic deviations between the assumptions underlying the perfectly competitive market model and what we have to cope with in the real world.[66] By introducing the concept of perfect competition in Sect. 6.3, we have already alluded to the fact that a market structure assuming that the individual firm does not have any control over the price of its product can only partly explain what is going on in real markets. Therefore, there has been considerable attention in microeconomics given to how the equilibria in markets with different structures, like monopoly and oligopoly, are constituted and what can be said about their social welfare properties. Another issue is that the perfectly competitive model requires that the agents acting in the market are very well informed, particularly about the quality of the products under consideration. In reality, however, it is very often much easier for the producers to find out about the properties of products they sell than it is for consumers to find out about what they buy. This *asymmetric* information distribution between consumers and

[65] This can be disentangled by an exercise in "geometrical accounting": by the consumption of the additional unit, consumer 2 gains a consumer surplus illustrated by the area $\bar{x}_2(\bar{x}_2 + 1)ed$. This is equal to the area $\bar{x}_2(\bar{x}_2 + 1)fd$ − the area $def$. Consumer 1, consuming one unit less loses consumer surplus illustrated by the area $(\bar{x}_1 - 1)\bar{x}_1ab$. This is equal to the area $(\bar{x}_1 - 1)\bar{x}_1ac$ + the area $abc$. So the change in total consumer surplus is illustrated by $\bar{x}_2(\bar{x}_2 + 1)fd - def - (\bar{x}_1 - 1)\bar{x}_1ac - abc$. Since the areas $\bar{x}_2(\bar{x}_2 + 1)fd$ and $(\bar{x}_1 - 1)\bar{x}_1ac$ are identical to each other, the total change in consumer surplus is $-def - abc$.

[66] Just one of the advantages of the perfectly competitive market model is that it provides a theoretically sound background against which many real world phenomena can be contrasted and analysed. Specifically, microeconomics investigates how equilibria and their social welfare properties are affected if we deviate from the assumptions underlying perfect competition.

producers has considerable consequences for the nature of market equilibria and their social welfare properties. These kinds of problems and how to cope with them is the focus of a special division of microeconomics: the economics of information.

Closer inspection reveals that market equilibria lose the amazing feature of being socially optimal if we deviate from the ideal market model sketched in the previous section, and allow for "distortions" such as market power and asymmetric information. Economists are often said to be brainless followers of the "free market ideology". Well, maybe some are (even though we hope it isn't true *at all*). Anyway, most of them are being strict judges: if markets do not meet the ambitious goal of producing socially optimal equilibria, then they are said to fail. *Market failure* is a key topic in microeconomics.

However important these and other issues may be for microeconomics and for the way markets operate in the real world, they are not at the core of our analysis. This is so because the concept of this book is to focus on problems directly related to the environment. Indeed, there is a deviation in real world markets from the framework of a perfectly competitive economy which is far more consequential for environmental problems, environmental policy, and the economic analysis of these issues than are the issues of market power and imperfect information. It is to this issue that we turn to for the remainder of this subsection: market failure generated by *externalities*.

The key to understanding this problem is that in the model of the perfectly competitive market, as sketched above, there are no agents other than consumers and firms. Moreover, all the relationships that exist between these agents are mediated by the market. Individuals interact with each other solely based on the price mechanism.

Obviously this stylisation is completely inappropriate if environmental issues are taken into consideration. Most of the effects that human activities have on the quality of air, water, soil, and biodiversity are not mediated by the market mechanism. Nevertheless, these natural resources are scarce. Consequently they should be at the core of economic analysis dealing with all forms of scarcity, be they recognized by the market mechanism or not.

This has important consequences for microeconomic analysis and for the social optimality of the market mechanism in particular. To see this we allow a third type of economic agent to enter the stage of microeconomic modelling, which has so far exclusively been occupied by consumers and firms. Somewhat ironically, we might call the third agent to be "the breather". Just as the traditional consumer, this agent wants to enjoy the consumption of a certain good but the object of his/her desire is not sold in the market place: Instead, it is clean air.[67]

[67] In reality, the breather is obviously not a third type of an agent. All individuals have to breathe, consumers and producers alike. So the different types of agents described in the economic model represent different roles that one and the same individual might take on. Of course, one individual playing more than one role, simultaneously, might also apply in the simple model that allows for consumers and firms only. All the individuals that act as producers are also consumers, although the opposite does not necessarily hold.

If the quality of the air deteriorates, the utility of the breather decreases. If the deterioration of air quality is a result of productive activity, the decrease in the utility must be counted as an opportunity cost of production. Using up clean air is a cost in economic terms, just as with using up any other scarce resource. Of course, to evaluate the worth of the clean air lost by production is more difficult than to evaluate the use of labour and capital. Regarding the latter, the prices for these productive factors in competitive markets are usually accepted as representing these values.

Regarding the utility individuals lose due to air quality deterioration, the same problems arise as those that were discussed when we dealt with the utility consumers derive from the consumption of commodities and services, as in Sect. 6.2, above. Consequently, the pragmatic idea to solve this problem is also the same: in microeconomics, the willingness to pay for clean air is taken to be a proxy for the utility people derive from clean air.[68] Please do not confuse this concept with the idea that people should have to pay for clean air. The former idea is not this kind (or any other kind) of policy recommendation, but a concept of evaluation. From the point of view of microeconomic analysis, it is a very welcome property of this concept of evaluation that it measures the utility from air quality (and the disutility from losing this quality) in monetary terms. This is a distinct advantage because the cost of reducing air quality can be added to other costs of production and can be therefore compared with the utility of production, which is also measured in monetary terms using the willingness to pay concept explained above.

Of course, what has been elaborated for the example of the utility loss the breather suffers if air quality deteriorates can be generalized to all environmental resources used up by human activity.

In general terms, we talk of *environmental damage* to denote all environmental resources which are used up (or the quality of which is reduced) by economic activity, irrespective whether what is affected is air, water, soil, or biodiversity. If we apply the willingness to pay concept to physical environmental damage, as briefly explained above, we arrive at a monetary number for environmental damage. When you read the expression "environmental damage" in economic texts, it usually refers to the monetary expression.

Just as has been done with the willingness to pay of consumers and the marketable resources used by the firms, microeconomics works with the assumption that the level of environmental damage relates to the level of economic activities and that this relationship can be expressed in terms of a mathematical function. In introductory economic essays like this one, these functions are presented in the simplified form of two dimensional graphs.

[68] An alternative measure would be the amount of money necessary to compensate the people for a certain loss in air quality. There has been a lot of discussion in the environmental economics literature about the relationship between these two evaluation methods: willingness to pay and the requirement to be compensated. See, e.g., Tisdell (2010). A particularly critical view is taken by Hahnel (2011).

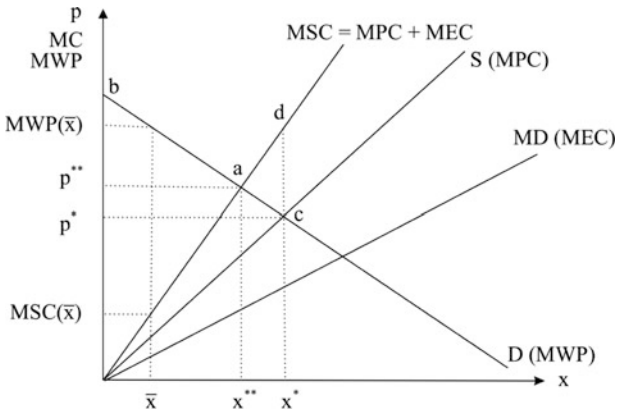


**Fig. 6.14** Market failure due to negative externalities

To make things most easy, for a start, we assume that production generates emissions proportionate to the level of output. This means that for any unit of the good $X$ produced, a constant amount, $a$, of a certain pollutant, $E$, is emitted. So the relationship is $e = ax$, where $e$is the level of pollution of type $E$ emitted, and $a$ is constant. These emissions generate environmental damage, $D$, depending on their level.[69] As has been explained for consumer willingness to pay and for the cost of production, we deal with environmental damage in marginal terms in microeconomic modelling. Ergo, we focus on how much damage is done by *an additional* unit of emission. In the simple model where emissions are proportionate to output, a statement on how much damage is done by an additional unit of emission is easily translated into a statement on how much damage is done by an additional unit of output. In analogy with the terminology used for willingness to pay and cost of production, additional damage is called *marginal damage*. Most often it is assumed that marginal damage increases with the level of emissions (the level of production): "it gets worse" with higher levels of economic activity. The reason for this assumption is a certain mathematical convenience on the one hand. On the other hand, it is ecologically quite plausible. For many pollutants the environment has a certain assimilative capacity. It can cope with a certain emissions burden quite well. The more this capacity is strained, however, the greater the damage to the assimilative capacity.

In Fig. 6.14 you find marginal damage integrated in the diagram illustrating a competitive market, as explained above. According to what has just been said, the marginal damage-curve (MD) is indicated as increasing with the level of production, $x$.

The inverse supply curve contains the marginal costs paid by the firms for the use of productive factors bought by them in input markets. The marginal damage curve includes the marginal cost of using environmental resources *not* bought in the

[69] We use the terms "emissions" and "pollution" as synonyms.

market in the production process. To terminologically distinguish these two kinds of marginal cost from each other, the marginal cost for the inputs bought in markets are called *marginal private cost* (MPC); the costs for the environmental resources not bought in private markets are called *marginal external cost* (MEC). They are "external" with regard to the market system. Society must bear, as a consequence of production, the sum of private and external costs. This sum is called "social cost". The social cost of an additional unit of production is called *marginal social cost*. In the graph, marginal social costs are illustrated by the MSC-curve. The MSC-curve is graphically constructed by adding up the MPC- and MEC-curves vertically.

The inverse demand curve contains the marginal willingness to pay of consumers for the product under consideration. Marginal willingness to pay is taken to be a measure for the benefits consumers gain from having a small additional unit of $X$.

We can use the figure to illustrate how the introduction of environmental issues (using the concept of external cost) affects the level of socially optimal output and the social optimality of the market equilibrium. The idea of socially optimal output has been explained to be one of maximizing the difference between the benefit of production yields for the consumers and the value of the resources used up in the process of production. This general concept does not change when we introduce environmental problems. However, the concept of resource use changes, in that the value of environmental resources has to be added to the value of resources bought in the input markets. Optimal output is defined as the maximization of the benefit to consumers minus social costs. Stating this in marginal terms, we find that the optimal quantity of production is defined by an equalization of the marginal willingness to pay and the marginal social cost. This condition is met at quantity $x^{**}$ in the figure. For any unit of output between 0 and $x^{**}$ marginal willingness to pay (represented by the respective point on the inverse demand curve) is higher than marginal cost (represented by the respective point on the inverse supply curve). This is shown for an arbitrarily chosen output level $\bar{x}$ in the figure. The contribution of production unit $\bar{x}$ to economic welfare is illustrated by the distance between the demand curve and the marginal social cost curve above $\bar{x}$, i.e., $MWP(\bar{x}) - MSC(\bar{x})$. Total welfare generated by the socially optimal production quantity $x^{**}$ is graphically illustrated by the area between the two curves within the limits of 0 and $\bar{x}$, $0ab$.

There is a fundamental divergence between the situation illustrated in Fig. 6.14 and the one illustrated in Fig. 6.11. In the previous model, where all the relationships between the economic actors are market relationships, the equilibrium quantity provided by the perfectly competitive market for $X$ is socially optimal, $x^* = x^{**}$. In the present analysis, allowing for environmental damage as an effect external to the market, the socially optimal level of production is smaller than the equilibrium level, $x^{**} < x^*$. The $(x^* - x^{**})$ units of X that are produced in the perfectly competitive equilibrium in excess of the socially optimal quantity can easily be seen as reducing social welfare. The value of scarce resources (whether being bought in the market or not) is higher than the benefit provided to its consumer, for each of these units. Therefore, according to the criterion of social

welfare maximization, these units should not be produced. This stated, they nonetheless are in the competitive equilibrium.

In this case, the market fails to produce a socially optimal equilibrium. Any situation where this occurs is called *market failure*. The reason for the market failure in the present context is that the environmental damage generated by production creates an *externality* (*external effect*), i.e., a consequence of economic activity not acknowledged by the market system.[70]

Diagnosing market failure and being unhappy about it does not mean that the market system as an allocative mechanism has to be thrown "onto the rubbish dump of history".

However, it is necessary (and intellectually attractive) to ponder what can be done to make amends. Specifically, it is here suggested that there is a role for the state (the government) to improve the situation. Microeconomists have been particularly interested in designing governmental interventions which improve the allocative results of the market system and do not hamper the socially beneficial allocative powers of this system. In furthering this idea, the question that pertains is: how do we design a smart system of governmental assistance for the *invisible hand* of the market? Answering this will be a subject of the next chapter. Obeying the concept of this book, we focus on environmental problems and the possibility of solving (or at least attenuating) them through sound environmental policies.

**Review Questions**

1. What are the general features of decision making of a baker, a manager of a basketball team, and a couple getting married?
2. Why is it necessary to choose between alternatives?
3. What might be the "opportunity costs" of reading this textbook (not to mention the benefits!)?
4. Which factors might influence an individual consumer's decision on how to divide his/her time between work, leisure, and education?
5. When deciding on the bundle of consumption goods to buy, what is the goal and what is the constraint of an individual consumer?
6. Can we expect that two consumers with identical budget, confronted with identical consumption goods and identical prices will choose an identical bundle of these goods?
7. What is the meaning and the purpose of a "comparative static analysis"?
8. Through what kind of thought experiment can a firm succeed in identifying its profit maximizing quantity of production?

[70] Externality is the most important cause of a market failure in the context of environmental economics. A related problem is generated by collective (public) goods, as referred to in subsection 2.3.3 and Chap. 4. However, please recall that we have pointed to other deviations from the "ideal" conception of a market system, which also lead to market failure, above. The most important ones are market power (particularly in the form of monopoly and oligopoly) as well as imperfect information (particularly in the forms of asymmetric information). These issues are dealt with in intermediate microeconomics textbooks like Varian (2010).

9. What is the meaning of "perfect competition"?
10. Starting from individual demand and supply decisions, how are "market demand" and "market supply" arrived at?

**Exercises**

1. Please find an example for a situation in which an individual consumer is forced to decide between alternatives!
2. Please elaborate the common features, as described in microeconomics, of the decisions made by
   (a) An electric power-supply company;
   (b) A cook;
   (c) A professional rugby player.
3. What are the functions of an allocation mechanism?
4. What is the difference between judging a market outcome as "good" or "bad", respectively, from an individual perspective, on the one hand, and a societal perspective, on the other?
5. Imagine that you have realized your individual equilibrium demand for pizza, on the one hand, and for burgers, on the other, given your income and the prices of both goods. Now assume that the price of pizza doubles, while the price of burgers remains constant, as does your income. How would you change your demand for pizza and burgers, respectively?

## References

Becker G (1968) Crime and punishment: an economic approach. J Polit Econ 76:169–217
Becker G (1993) Nobel lecture: the economic way of looking at behavior. J Polit Econ 101:395–409
Blaug M, Lloyd P (eds) (2010) Famous figures and diagrams in economics. Edward Elgar, Cheltenham
Cooter R, Ulen T (2004) Law and economics, 4th edn. Pearson, Boston
Eaton BC, Eaton DF, Allen DW (2011) Microeconomics: theory with applications, 8th edn. Pearson, Toronto
Estrin S, Laidler D, Dietrich M (2008) Microeconomics, 5th edn. Pearson, Harlow
Hahnel R (2011) Green economics. M.E. Sharpe, Armonk/London
Perloff JM (2007) Microeconomics, 4th edn. Pearson, Boston
Rawls J (1971) A theory of justice. Harvard University Press, Cambridge, MA
Tisdell CA (2010) Resource and environmental economics. Modern issues and applications. World Scientific Printers, Singapore
Varian HR (2010) Intermediate microeconomics. A modern approach, 8th edn. Norton, New York