The Theory of Demand

11

In Chap. 4 we saw that, according to the usual economic worldview, the sole purpose of economic life is to produce goods for purchase by consumers. Producing the right goods in the right amounts, with the characteristics consumers desire, is what an efficient economy should be doing as much of the time as possible. Clearly, in order to translate this broad objective into specific policies we require a theory of the consumers themselves: what governs the choices they make and how their individual decisions in the marketplace affect their ultimate well-being.

As we will see, however, economics has developed a theory of consumer choice that is nearly useless for these purposes. (This is not a controversial statement.) It is very elaborate and contains more than a few valuable insights, but it falls far short of what policy-makers, or marketers for that matter, are looking for. Yet this is not entirely fair, since the purpose of the conventional economic theory of consumer choice is not to answer substantive questions about the impact of consumption on well-being or to predict future consumption patterns, but simply to identify the conditions on the demand side of the market that must be met for the Market Welfare Model to hold. In other words, its purposes are internal to economic theory itself, rather than outward-looking or pragmatic.

This is an important function, one we will take seriously in the pages to come. Nevertheless, if we were to end the story at that point many readers would feel frustrated. They may be interested in the logical nuances of normative economic modeling, but they also want to know whether the economy they are living in is truly delivering the goods, as measured by the well-being of the population. An economics text is not the place for detailed examination of such questions, or even of the theoretical tools such an examination might employ, but we will survey briefly some of the main ideas that have emerged in two alternative approaches to consumption. The purpose is to set in relief what is truly unique about standard economic theory in this field.

P. Dorman, *Microeconomics*, Springer Texts in Business and Economics,
DOI 10.1007/978-3-642-37434-0_11, © Springer-Verlag Berlin Heidelberg 2014

11.1 Utility and Utilitarianism

The place to begin is language, specifically the central term in economic discourse about consumption, **utility**. We have been using this word loosely up to this point, but it is easily misunderstood and deserves further clarification. As mentioned previously, what utility does *not* mean in economics is usefulness. A completely useless item can still give people utility if they desire it for some reason. Sometimes a hot fudge sundae can offer more utility than a healthy, nourishing salad.

What utility does mean is difficult to say precisely. We imagine that people could sum up their happiness in a given situation with a single evaluation and then say whether, on balance, they were better off than they would be in another situation. If the difference between the two situations is that in one an individual has less money but a hot fudge sundae, and in another she has more money but no sundae, the comparison tells us whether the extra money is worth the gooey pleasures of the sundae. Economists would say that the utility of one is greater or less than the utility of the other. In other words, utility is the measurement of desire corresponding to the preference of one thing, or group of things, over another. If I want ice cream rather than salad, and I am willing to actually make this choice if given the opportunity, this means that, in this situation, I get more utility from ice cream than from salad.

We can imagine that people might make every possible comparison between different assortments of goods, with and without different amounts of money, and at the end of the process produce a complete set of rankings. For each comparison they are prepared to say which they would prefer, or whether they evaluate both exactly equally. From this we would be able to rank all the possible choices from highest to lowest, and this would also tell us which choices were above the others in terms of utility. Economists call such a ranking a "preference map" and they see it as the best guide to the study of consumer satisfaction.

In this book we are going to cut a few corners. Rather than make the minimal assumption that individuals do no more than compare sets of options (but this is already a lot, since they have to compare *all* such sets), we will go further and assume that they can actually put a numerical measure on each choice. Like Olympic judges, they give this sundae an 8.7 and that salad an 8.3. Doing this for every possible good produces a complete numeric scale, with a utility score for each. Since this entails more demanding assumptions, economists will be a bit uncomfortable with it, but nothing of importance for our survey will be lost; nearly everything we will be doing with this numerical conception of utility also works, but in a more complicated way, with the just-make-comparisons approach.[1]

It will be useful to step back for a moment and consider the implications of the analysis we are about to embark on. People are making choices in the market; they are paying money and buying goods and services. This is observable and even measurable: willingness-to-pay, after all, can actually be computed from economic

[1] In the language of economics, we will use a cardinal rather than ordinal approach to utility.

data. What we would like to know is how all this buying affects people's true well-being. This is invisible and possibly unmeasurable. The problem is to infer the second from the first, if possible. Economists typically make the assumption that at the individual level the two correspond to one another perfectly, that if any good is chosen over any other it makes the individual better off as well, and that willingness to pay is a satisfactory numerical measure of how much additional well-being a consumer can expect to get from an item he purchases.

All of this is incorporated into the concept of utility. Utility is the element of well-being corresponding to the units of money people spend on things. In such a scheme people are never disappointed; the goods they buy deliver exactly the payoffs they anticipate, which in turn are encoded in the prices they are willing to pay. As a theory it is difficult to justify, but it has the convenience of enabling economists to discuss well-being (normative economics) with exactly the same tools they use to analyze observable consumption behavior (positive economics).

Before we dismiss the whole enterprise as improbable, we should consider the case in its favor. It rests primarily on the question, if you don't trust an individual to make the choice that will turn out best for her, who do you trust? If the expectations of utility people have in their minds when they make their purchases differ in some way from the well-being they actually receive, does it matter if this is still the best guess anyone can make about what the effects will be? In that case we could say that utility theory is approximately correct, and that following its guidelines is the best course of action. It's a bit like saying you have a thermometer that sometimes gives too high a temperature and sometimes too low. You can acknowledge this, but if this is the only or best instrument you have, and if you are unable to tell what the error is for any particular reading, all you can do is record the temperature it gives you and hope for the best. This argument, however, depends on the strong claim that there is, in fact, no better guide to human well-being than consumer willingness-to-pay, and, as we will see, there are many who would disagree.

There is also a political aspect to the question posed in the previous paragraph. Surely individuals deserve a benefit of the doubt in their choices on the grounds that this safeguards their autonomy to choose as they please. The danger in having some other theory of well-being is that it can justify intrusions by well-meaning authorities that put individual freedom at risk, the problem of **paternalism**. Logically, there is no requirement that an objective theory of well-being (one that can be determined by "outsiders" like academic researchers rather than the individual whose well-being is at stake) must necessarily lead to infringement of freedom, but there will typically be a temptation. Take the case of smoking cigarettes. This has severe health consequences but also provides at least some pleasure for smokers. We could let smokers decide for themselves whether the health cost is worth it. On the other hand, public health experts might conclude that, no matter what smokers may think, smoking makes them worse off. They could issue this opinion and leave it at that, but some would see it as a basis for laws restricting the freedom to smoke. (Taxes on cigarettes, which exist nearly everywhere, do this, for instance.) Is this a bad thing? Without passing judgment, it should be clear that, if nothing else, the lost freedom of smokers ought to be a consideration—it is not

without some value. For many economists and others who subscribe to the tenets of political liberalism (as understood in the context of this book), utility theory ("the smoker always makes the choice that maximizes his utility") is a bulwark against those who would give individual freedom too little weight.

One final point: utility theory performs the magic of making possible a reconciliation between liberalism and utilitarianism. Liberalism says that people should be free to make their own choices over how to conduct their lives, including what to buy in the market. Utilitarianism says that the goal of policy should be to maximize the total well-being of the individuals who make up society. They can coexist, even potentially, only if individual choice always serves to maximize individual well-being. If this were not the case, we would have to choose between choices that are free and choices that make people better off. Economists want to preserve both of these, and the simplest way to do it is to simply assume that they are compatible. Even so, however, we will see that the conditions under which free individual choice maximizes well-being are quite limited. (We should be prepared for this result after our repeated encounters with the Prisoner's Dilemma.) Thus there is something important to be learned from standard utility theory: we can make the most favorable possible assumptions about the relationship between free markets and human happiness, and even then we may find that the two diverge. That is why the subject is interesting and important.

11.2 Utility and Individual Choice

So now let us suppose that an individual whose utility broadly conforms to the description we've just considered is buying shoes. If he has no shoes at all, he probably has a strong need for a pair; that is, this first pair of shoes will give him a lot of utility. Perhaps he think he needs to stock up, however, and buy different shoes for different occasions. In what follows, to make things as simple as possible, we will assume that all shoes, whatever their make or purpose, cost exactly the same. If the first pair is for work, maybe he needs another pair for dancing. This second pair also provides plenty of utility, although not as quite as much as the first. (If it had provided more, it would have been the first pair he bought.) Still, there are other reasons to buy shoes: for walking in town, for walking in the mountains, for wearing to formal occasions and so on. Our shopper goes from one part of the shoe store to another, buying pair after pair. We can assume that each successive pair gives him a bit less utility. Finally he gets to the point at which he would not accept another pair of shoes even if it were given to him.

The situation is illustrated in Fig. 11.1, where the utility received from each pair of shoes is measured on the vertical axis and the number of pairs is measured on the horizontal axis.

Recall from Chap. 4 the concept of "marginal"; it refers to the additional amount of some quality. There we introduced marginal cost and marginal benefit, and here we will use the term **marginal utility**. Marginal utility is the additional utility someone gets from acquiring or consuming one more unit of a particular good.

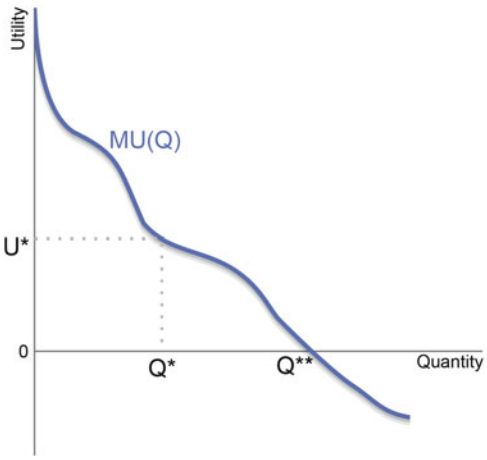

Fig. 11.1 Diminishing marginal utility from buying shoes. The U axis measures the marginal utility an individual receives from buying a pair of shoes; it is negative below 0. The Q axis measures the number of shoes purchased. Marginal utility declines until it reaches 0 at Q** pairs. If U* is the utility corresponding to the price of shoes (assumed constant), Q* is the number that will be bought

In this example it is the extra utility that comes from buying an additional pair of shoes. I have drawn the marginal utility curve MU(Q), which represents marginal utility as a function of the number of shoes being purchased, as downward-sloping. What this says is that, the more shoes a person buys, the less additional utility he will get from each additional pair. Why assume this? Economists suspect that this pattern holds for the vast majority of goods and services people acquire and call it the "law" of diminishing marginal utility.

(When Ferdinand Marcos, president of the Philippines, was forced to leave office after popular demonstrations in 1986, it was discovered that his wife Imelda had amassed a collection of 3,000 shoes. What surprised the public was not her wealth, which was well-known, but the implication that the law of diminishing marginal utility did not seem to apply to her, at least in the realm of footwear.)

To read Fig. 11.1, begin at the left of the diagram, at the very first pair of shoes bought. Here the MU curve is at its highest, indicating that this first pair is strongly desired. As we move to the right along the Q axis, we are observing the second pair purchased, then the third and so on. The downward slope of MU indicates that the marginal utility of each subsequent pair is declining. At Q** (greater than 3,000 for Imelda) the curve enters negative utility territory, signifying that, even if money were no object, the individual would stop acquiring shoes; they are more trouble than they are worth.

Of course, money *is* an object. We will continue to suppose that all the shoes sell for the same price, and that there is a utility corresponding to that price that we can designate as U*. For example, if a pair costs $40 U* represents the utility of having an extra $40 in your pocket. As long as the utility acquired from an additional pair of shoes exceeds U*, it makes sense to buy it. In our diagram this is true for several pairs. If the utility of the money is greater than that of the shoes, however, no further purchases will be made. Thus the utility-maximizing shopper will stop at Q* pairs. This last pair just barely justifies itself, and any more would not be worth the price.

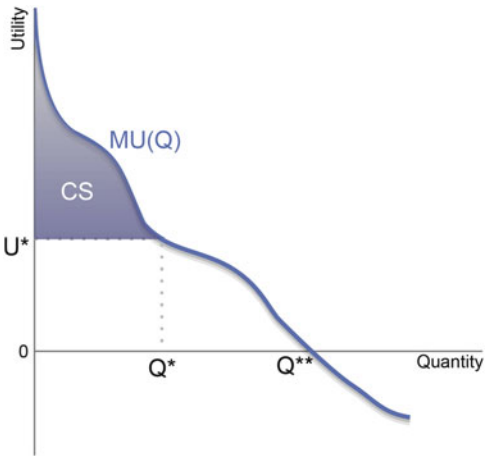

Fig. 11.2 Consumer surplus, measured in utility, gained from buying shoes. This diagram is identical to Fig. 11.1, but with the addition of consumer surplus, the difference between the utility gained from a pair of shoes and given up due to the money paid for it, summed over all the pairs purchased

We can imagine what the impact would be of a change in prices. If the price of shoes goes up, for instance, so will U*: more money translates into more utility. The horizontal line at U* will intersect the MU curve at a lower number of shoes; Q* will go down. This is exactly what we would expect, of course, and it shows that the utility story we are telling is consistent, at least in this respect, with common sense.

What is happening to the well-being of the person buying all these shoes? If the utility equivalent of the price is U* and the quantity he is purchasing is Q*, the last pair leaves him no better or worse off, but every other pair gives him a net addition to his "store of utility". If Q* is five pairs, for example, then the marginal utility of each of the first four pairs exceeds the utility given up to buy them, U*. Figure 11.2 illustrates this. It is identical to Fig. 11.1, with the addition of a shaded area representing the net increase in our hero's welfare—the sum of the net utility gains resulting from all pairs up to Q*. This is referred to as the consumer surplus, here measured in units of utility. What determines its size are three factors, the slope of the MU curve, the level of Q* and the level of U*. A steeper slope, more Q* and less U* all contribute to greater consumer surplus.

If we think back to Chap. 4 and the claim that the sole purpose of having an economy is to increase consumer welfare, consumer surplus is the key to it all. It is not the utility given up by spending money that measures economic success, nor the total utility gained from the goods purchased, but the second minus the first, at least for this one person and this one commodity. More consumer surplus signifies greater economic gain.

This is all well and good but, unfortunately, utility, as we have seen, is invisible and unmeasurable (if indeed it is a meaningful concept at all). What can be observed is not utility but money. So let us look at the same situation in money terms, as in Fig. 11.3. It is identical to Fig. 11.2, except that, instead of utility being measured on the vertical axis, it is money. Instead of a marginal utility curve, we picture a demand curve whose height at any particular Q is the consumer's **willingness to pay**. We can directly observe the price actually paid, P*, and we

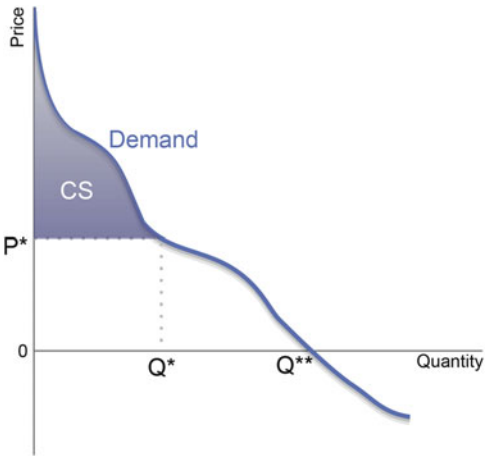

Fig. 11.3 An individual demand curve and consumer surplus for buying shoes. This diagram is identical to Fig. 11.2, but expressed in terms of money rather than utility. D is the demand curve for an individual, P* is the price charged, and Q* is the amount purchased at that price. Consumer surplus is represented by difference between willingness to pay (the height of the demand curve) and the price, summed over all the goods purchased

can, in principle, ask the consumer how much he would be willing to pay for every pair of shoes, from the first one he buys to those he would not buy at the current price. (Some of the readers of this book may have been asked exactly this sort of question by market researchers at shopping centers or other public places.)

Figure 11.3 is real in a sense that Fig. 11.2 is not. Prices are real, and so is willingness to pay. Utility is imaginary, an idea conceived by economists and philosophers but not directly measurable in the way that prices are. Nevertheless, from the standpoint of normative economics (how to make people better off), it is utility—Fig. 11.2—that matters, not money. The question naturally arises, what exactly do we need to infer Figs. 11.2 from 11.3? The answer, aside from the whole apparatus of utility itself (which we discussed at the beginning of this chapter), is what we might call the "exchange rate" between money and utility. That is, for any given amount of money in Fig. 11.3, what is the corresponding amount of utility in Fig. 11.2? The exchange rate analogy is helpful; you could think of these two diagrams as representing the same thing but in different currencies, like euros and yen. So many euros are worth so many yen, and similarly for money (in any currency) and utility. The name given to this exchange rate by economists is the **marginal utility of money**. Like a currency converter, it tells you how many units of utility an individual gets per additional unit of money and vice versa. If we can believe that something like this exists in the mind of our hypothetical shoe-buyer, we can go back and forth between diagrams 2 and 3 without great difficulty.

11.3 Market Demand, Consumer Surplus and Utility

The next step is to bring all the consumers together and examine the demand for shoes throughout a given market. (This market might be local, national or global depending on the purposes behind our analysis.) To see the relationship between

Table 11.1 Number of shoes purchased by three consumers at various prices

Price per pair	Huey	Dewey	Louie	Total
$20	8	5	3	16
$30	8	4	1	13
$40	8	4	0	12
$50	6	2	0	8
$60	5	1	0	6

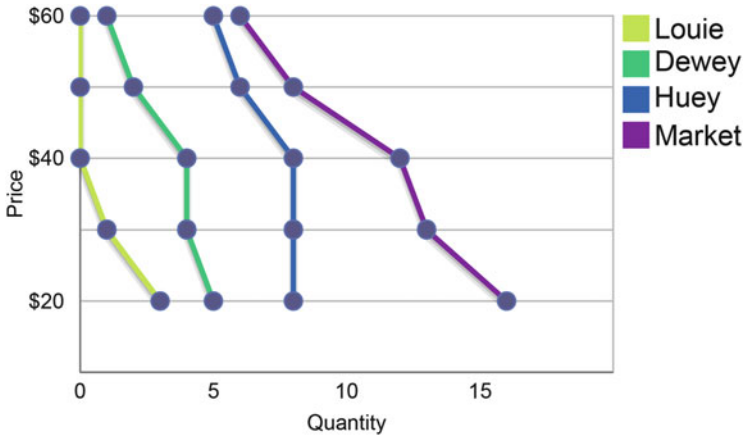

Fig. 11.4 Individual and market demand for shoes. Individual demand curves are given for three consumers and for the market consisting of all three. The market demand curve is the horizontal sum of the three individual demand curves

individual and market demand, consider the hypothetical **demand schedules** of three consumers, Huey, Dewey and Louie.

Table 11.1 tells how many pairs of shoes each is willing to buy as the price rises from a low of $20 to a high of $60 per pair. These data are plotted in Fig. 11.4.

This diagram demonstrates the relationship between the individual demand curves and the market demand, when the market consists of just these three. At $60, for instance, Huey buys 5 pairs, Dewey 1 and Louie none, so the total is 6. Tracing this for each of the possible prices constructs the market demand curve as the horizontal sum of the individual curves. As long as no individual will buy *more* at a higher price (as long as individual demand curves are either vertical or downward-sloping), the market demand curve will never be upward-sloping. The negative relationship between the market price and the amount consumers want to buy is called the **law of demand**. Like all laws it is sometimes broken, but it holds in the vast majority of cases.

The second implication is that each point on the market demand curve represents someone's willingness to pay for that item. When the price falls from $40 to $30, for instance, the market demand goes up by one. That "one" is Louie, who buys his first pair of shoes at that point. He is willing to pay $30 but not $40, so his willingness to pay is represented by the market price. (Because of the large price

intervals, he may be willing to pay more, but let us assume this represents the most he would pay. If space were not a constraint in this book, we could watch the price fall penny by penny.) We can call Louie the **marginal consumer**, the individual whose preferences are represented by the point on the demand curve corresponding to $30. Thus, every point on the demand curve "belongs" to a marginal consumer somewhere and represents his or her willingness to pay.

A third implication is that we can sum the consumer surpluses of the individual consumers in order to calculate consumer surplus for the market as a whole. Suppose the actual price charged turns out to be $40 per pair. Total demand will be 12 pairs, of which Huey will buy 8 and Dewey 4. (Louie has been completely priced out.) Of the 8 pairs purchased by Huey, he would have bought 5 at $60 and a sixth if the price were to fall to $50. The final two he buys only when the price falls further to $40. This means that five pairs give him a consumer surplus of at least $20 each and one pair at least $10. How large the surplus is we cannot say, since we don't have information on intermediate price levels, only on ten-dollar increments. For instance, perhaps one of the final two would have been purchased at $45 dollars rather than $40; this mean it would add another $5 to his consumer surplus. We do know the minimum, however: it is $110. For Dewey this same amount comes to at least $30. Now turn to the market demand, which is 12 pairs when the price is $40. Of these, six are worth at least $20 more than that to their buyers, because they would be bought at $60, and another two are worth $10 more. Thus the market consumer surplus is at least $140, which is the sum of the two individual surpluses.

From this simple exercise we can see how individual demands sum up to the market demand, but what about utility? We were able to go from Figs. 11.3 to 11.2 with a few handy assumptions; is there any way to translate Fig. 11.4 from money into utility units?

Recall that the key to translating money into utility at the individual level is the marginal utility of money. The problem at the social (market) level is that each person is likely to have a different exchange rate. There are two general reasons for this. First, some people are more materialistic than others. Henry David Thoreau and Mohandas K. Gandhi were both famous for placing other values above material ones; they could be said to have had low marginal utilities of money. Others have an insatiable craving for things that money can buy; their marginal utilities will be higher. The second reason is that money is likely to obey the law of diminishing marginal utility in the same way most other goods do. The more money you have, the less additional utility you get from an additional dollar. Equal dollar amounts have very different utility significance for rich and poor (See Box).

Box 11.1: Traffic Fines in Finland

In most countries fines for violating the law are set in monetary terms. A parking ticket is a certain sum of money no matter who has to pay it. This is fair in some respects, but it puts a greater burden on low-income groups.

(continued)

Box 11.1 (continued)
Wealthy people can ignore fines that would create a small crisis for someone living on a tight budget. Finland is different, however. Finland sets fines as a percentage of the violator's income in order to equalize the utility cost paid by offenders. Other European countries do this too, although Finland is unique in having no ceiling on the amount that can be assessed.

On a June day in 2000, police in Helsinki pulled over Anssi Vanjoki for doing 45 miles per hour on his motorcycle in a 30 mile-an-hour zone. Because Vanjoki was a senior vice president for the cell phone company Nokia and had earned over $5 million the previous year, his ticket came to $103,000. Vanjoki appealed, arguing that his income had suffered a nosedive in 2000 and that police should have taken it into account. He won, and the fine was reduce to "only" $5,245. Other wealthy Finns have been fined in the tens of thousands of dollars for comparable offenses.

The Vanjoki case set off a debate in parliament. Some legislators argued for scrapping the system and setting fixed monetary amounts for small offenses, but not all. Parliamentarian Annika Lapintie was quoted as saying, "The law is a deterrent. It would be totally unjust if the poor and wealthy pay the same because the wealthy wouldn't feel it." In between were lawmakers searching for a compromise, keeping the percentage of income formula, but putting a cap on it to avoid potentially embarrassing outcomes.

What this means is that, in general, it is not possible to infer utility amounts from dollar amounts at the market level. For instance, suppose that consumer surplus in one market is $10,000 and in another it is $15,000. We can't conclude that the surplus in utility terms is greater in the second than the first, because it is possible that the average marginal utility of money in the second market is substantially less.

Naturally, economists find this state of affairs frustrating. They want to be able to make judgments about which policies will make people better off, but they are lacking a crucial piece of information they would need to convert monetary measurements into assessments of human welfare, since marginal utilities of money are unobservable and nearly impossible to estimate. In the end, they have these options:

- They can assume that the average marginal utility of money in a group is a function of its average income (which can be measured). One way to do this would be to express monetary values as a percentage of income rather than an absolute amount; this is a strategy similar to that used in Finland in Box 11.1. This approach assumes that differences in income are primarily responsible for different utility values of money, or at least that the differences due to personal values will mostly cancel out at the group level.
- They can assume that all marginal utilities of money are the same. If the groups are relatively similar in composition this may not be too much of a stretch. In fact, often the comparison is between different consumer surpluses for the same

market, when different policies are being considered. The specific people whose surpluses are being summed may change somewhat from one policy to another, but often not greatly. Another argument is that, if we have to make many decisions that will affect consumer surplus in a wide variety of markets, differences in the marginal utility of money will largely cancel out. For instance, it is not likely that the groups that benefit from a particular bridge being built, or from lower postal rates for certain types of magazines, or from publicly financed research into specific diseases will all be disproportionately rich or poor, even though any one such group might be. Thus, over the course of a large number of economic policy decisions, the goal of maximizing consumer surplus may yield results that are in the interest of all social groups. On the other hand, it might also be the case that, for many such decisions, the richest and poorest citizens may indeed find themselves lined up largely on opposite sides.

- They can refrain from making utility comparisons at all. This approach is theoretically unimpeachable; if you would need to know everyone's marginal utility of money in order to say which consumer surplus corresponds to the most utility, and if there is no way to get this information, why not just give up? The problem (or challenge) with this choice is that it greatly limits the number of decisions that can be justified with economic analysis. (We will look at the consequences of this approach more closely in Chap. 21.) Yet it is not always possible to pick and choose between decisions; often they simply have to be made on *some* basis. If not judgments of utility, what? We will return to this question later in the chapter.

We are now in a position to sum up the significance of the utility-based theory of demand for the Market Welfare Model. Recall that the model puts forward three premises and draws one conclusion:

Market Welfare Model

Conditions

1. The demand curve represents the marginal benefit to society from the consumption of some good.
2. The supply curve represents the marginal cost to society from the production of this good.
3. The supply and demand curves have a single, stable equilibrium.

Conclusion

The market equilibrium maximizes the net benefits to society of the production and consumption of this good.

The utility theory of consumption puts the first premise under a microscope. It identifies the underlying conditions which, if all met, would enable the premise to be accepted. For convenience, they are listed in Table 11.2.

If all of these things are true, logically the first premise of the Market Welfare Model follows as well. In our brief survey of utility theory, we assumed the first item in this table for the sake of discussion. The second and third were results of the analysis and, at least for now, appear plausible. The fourth point is dubious but not impossible (and there is always the hypothetical transfer of money to set things

Table 11.2 Sufficient conditions for the demand curve to represent the marginal benefits to society

1. Benefit to society can be represented as the sum of individual utilities
2. Each point on the market demand curve represents the willingness to pay of the consumer who is just induced to buy this one item
3. That willingness to pay reflects the amount of utility of this marginal consumer
4. The ratio of marginal utility to willingness to pay, i.e. the marginal utility of money, is equal across consumers
5. There are no other impacts of the consumption of this good other than what is represented by consumer willingness to pay

straight). The final point will be discussed in Chap. 15; for now, let's also assume that it holds. The result is that much depends on the fourth point.

One interpretation of this discussion is that the Market Welfare Model is more plausible in societies with more equal income distributions, or with more active income redistribution programs. Free market allocation of a scarce but crucial good, like potable drinking water in some countries, has greater justification when there aren't large differences in income and wealth. To say that you would distribute water according to market principles is to say that those who are willing to pay the market price would get water and those who aren't won't. If these differences in willingness to pay reflect true differences in the need people have for water, there may not be a problem. If they reflect mainly differences in income, a too-rigid adherence to market principles may result in disease and death as the poor are cut off from their water supply. The preceding analysis of marginal utility of money is not academic hair-splitting; it is the basis for many present-day controversies in economic policy.

A second way to make this point is to recall the geometric argument for the Market Welfare Model, Fig. 6.6 from Chap. 6. This is reproduced below with one additional element as Fig. 11.5 on the following page.

Here we add consumer surplus to the picture. It comprises one portion of net benefits, the part acquired by consumers when the value of what they buy, measured by their willingness to pay, exceeds the price they have to pay for it. The marginal consumer at Q* gets no surplus, but all those to the left of her, who would be willing to buy at prices higher than P*, do receive a surplus. The area marked CS is the sum of all these individual surpluses. From this diagram it is clear that consumer surplus represents the consumer side of net benefits (as the area below P* represents the producer side). Unfortunately, we also recall that there is no necessary relationship between the size of consumer surplus in monetary and utility terms. Thus, insofar as the amount of net benefit depends on the amount of consumer surplus, the Market Welfare Model, if it is to be a model of welfare and not just money, requires some solution to the problem posed by different marginal utilities of money. Since we live in an imperfect world, we might say that the problem is not too great and get on with other tasks, but how great is it? The take-home message from this analysis is that the answer depends on the circumstances, so it pays to be informed.

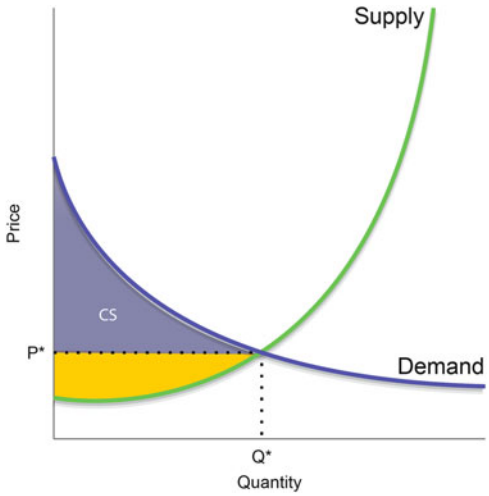

Fig. 11.5 Maximum net benefit and consumer surplus in the Market Welfare Model. The total *shaded area* represents total net benefit according to the Market Welfare Model, assuming the necessary conditions are met. The more heavily shaded area below the demand curve but above the price is that portion of the net benefit captured by consumers, the consumer surplus

11.4 Do Consumers Maximize Utility?

Let's continue to adopt the utility model, at least provisionally, but put aside the assumption that people are always rational—that they choose the option that maximizes their expected utility. If that is the case, there may be a gap between willingness to pay and utility due to systematic bias. In other words, people may regularly overestimate or underestimate how much utility they will get as a result of a purchase. Regular errors are those which are describable and predictable; is this true of consumer choice?

A very large body of research says it is. One of the main currents in behavioral economics is the study of consumer choice; it joins a more venerable effort by marketing specialists to figure out what makes consumers tick. Some of the biases that have been discovered are these:

- Faulty self-perception. People systemically overrate themselves in most respects: they think they are better drivers than they are, they plan to exercise more regularly than they actually will, and they buy ingredients for elaborate meals they will never make. While this affects many aspects of life, it has a definite bearing on consumption choices. They buy products for the person they think they are, not the one they actually are.
- Mental accounting. People divide choices into different categories and then make decisions separately for each category. If a restaurant meal is eaten on vacation, it may be assigned to the "vacation" category, and spending decisions will be made that would never occur at a similar restaurant under ordinary circumstances. Here's another example. Suppose you get money from a relative for your birthday; you might save part or all of it. On the other hand, what if you get a camera as a gift, but you already own one? You take it to a store and get money back for it; will you spend this money the same way you would spend the

money you are given directly? Laboratory evidence suggests you wouldn't: having already had a consumption good in the form of a camera, you place the money in a (mental) "consumption" account and spend it. The gift money, which goes into a different mental account, is more likely to be saved. Accounting heuristics like this lead to inconsistent behavior by consumers.

- Status quo bias. People are more reluctant to part with something they have than they are eager to acquire the same thing if they don't have it. In other words, exactly the same item will have two different prices for the same person, the price they would sell it at and the price they would pay. If the good in question is a small part of the person's overall wealth (and therefore has little effect on the marginal utility of money), there is no evident reason for this discrepancy other than a preference for what one already has.
- Misperception of risk. Many choices in life involve risks. We can buy a new car at a higher cost but a lower risk of repairs over the first few years or a cheaper used car that might turn out to be a money sink. The decision to borrow money, or to lend it, is risky, and so is choosing a specialized major in college that may not lead to a job. If people are to make such choices effectively, and not squander utility in a predictable way, they need to estimate risks accurately. Much research has shown, however, that this is frequently not the case. People place too much importance on very small risks of catastrophe and not enough on much more likely risks of moderate loss. They are unduly swayed by vivid examples rather than evidence of riskiness drawn from extensive experience. Partly because they overestimate their own abilities, they also give insufficient weight to risks that they think they may be able to influence, compared to those over which they have no control at all. These and other biases interfere with choices that will actually make people better off rather than play to their insecurities or, paradoxically, their sense of invulnerability.
- Poor forecasting of feelings. When considering a purchase, a consumer is often confronted by a strong anticipation of pleasure. If the item is something she wants, she will feel a surge of excitement at the thought of buying it. Research, however, suggests she is likely to attribute too much importance to these passing emotions. The long-term effects of positive economic events (like a purchase or an increase in income) on well-being are usually less than we anticipate, while this is less likely to be the case for noneconomic events (like a change in health or marriage). Especially for the consumption decisions economists are most interested in, then, people have a tendency to confuse the immediate emotional impact with the long-term effect on well-being, if any, that will remain after the initial jolt has worn off. The problem is made worse by the tendency, substantiated by psychological research, that the emotional state of the consumer at the moment a decision is being made has a significant impact on choice. (Don't go shopping for food when you're hungry.) If people were truly rational they would look past their transitory mood and consider their feelings down the road.

These and other traits represent systematic, rather than random, deviations from the postulates of rationality that economists have historically applied to consumers and other decision-makers. Taken together, they indicate that the choices people

make give them less utility than they might otherwise be able to get. Thus, if utility remains the guiding framework, we must make a choice between deferring to consumer behavior in the marketplace or devising policies to offset the loss of utility from faulty decision-making. But utility is not the only framework.

11.5 The Pursuit of Happiness

If it is really well-being we are interested in, why don't we just ask people how well off they are? This may seem too direct an approach to work; perhaps people wouldn't know, or maybe their answers would mean different things to different respondents. As it happens, however, researchers have been using survey methods to find out how happy or satisfied individuals say they are, and their results have been repeatedly validated. Thus, those who say they are happy with their lives are more likely to be described by others who know them as happy; they are more likely to smile and initiate social contact; they tend to live longer; they put in more effort and take fewer absences at work. Recently neurophysiologists have begun to compare the brain activity of happy and unhappy people, and they are discovering the patterns their training has led them to expect: neurotransmitters in the appropriate regions of the brain are fired in ways that correspond to subjective reports of happiness or well-being. All evidence suggests that we should believe the answers people give to questions about their emotional state.

The direct measurement of happiness promises to rectify two problems with the utility theory of consumption we have just reviewed. First, utility theory rests on a set of assumptions that may simply be wrong, particularly in believing that consumers are rational, reliable utility maximizers. Second, utility theory, based as it is on inferences from consumption behavior in the marketplace, applies only to goods that are traded in markets. Shoes can be assessed for the utility they offer but not the pleasure of seeing a rare bird in your backyard. Economists have proved clever at using market transactions to infer nonmarket prices, as we will see in Chap. 15, but there are limits to such ingenuity. Using survey methods to assess happiness has the potential to measure *any* factor that the researcher wants to find out about.

This second point is particularly important, since one of the large questions before industrialized societies is where to draw the line between economic and noneconomic activities. Should we spend less time working, even if this means producing and consuming fewer goods? Should we sacrifice economic growth to other considerations, like a greater emphasis on family life or a healthier environment? Utility theory can tell us something (maybe) about how economic goods should be traded off against one another, but not about the tradeoff between things we acquire through the economy and the aspects of life that are crowded out by working and spending.

How do economists and others do happiness research? An example is presented in Box 11.2, which explains how two economists put a value on being sexually active.

Box 11.2: Money Can't Buy Me Love

What is the contribution of sex to happiness compared to that of money? While it surely varies from one individual to another, an average relationship was worked out by two economists, David Blanchflower and Andrew Oswald, drawing on the data in a survey of 16,000 US adults.

To do this, they constructed a formula for predicting the answer people would give to a question about their overall well-being. If H is the number people report on a happiness scale, where a higher value of H signifies greater happiness, the formula would look like this:

$$\begin{aligned} &a + b_1{}^{*}\text{age} + b_2{}^{*}\text{gender} + \ldots + b_{n-1}{}^{*}\text{income} \\ &\quad + b_n{}^{*}\text{frequency of sex} + e = H \end{aligned} \tag{11.1}$$

Here a is referred to as a **constant** (I will explain it shortly), and the various b's are **coefficients**. The formula says that an individual's happiness score is approximately equal to a fixed number (the same for everyone) plus their age multiplied by its weight in the formula (the coefficient b_1) plus a number representing gender (say, 1 if male, 0 if female) times its weight b_2 + many more factors times their weights (signified by the "…") plus income multiplied by its weight b_{n-1} plus the frequency of sex times its weight b_n. If we plug in the values for any given individual for all n factors and apply the right values for a and all n weights (coefficients), the formula will tell us what his or her happiness score is likely to be. The formula won't be perfect, however. There will be error in its prediction, which is what e signifies. Finally, we can understand the meaning of the constant a by supposing that the value of every factor included in the formula is zero; then H would approximately equal a.

So where do all these numbers come from? The values for all the individual characteristics and the happiness score come from the survey. The values for a and the various b's come from statistical techniques designed to select them in such a way that this formula, applied to everyone in the survey, produces as little error as possible.

The weights are the whole point of the exercise. The sign of a given weight tells us whether the variable it is attached to makes a positive or negative contribution to H. For instance, suppose in the gender variable that 1 = male and 0 = female. If b_2, the coefficient (weight) attached to gender is negative, it indicates that, considered independently of all the other factors being studied, being male lowers the predicted value on a respondent's H score. In addition to the direction—positive or negative—of the weight, its size and significance matter too. The size tells us how big the effect is: if $b_2 = -0.1$, it says that being male, considered separately from everything else, lowers a person's predicted happiness score by a tenth of a point relative to females.

(continued)

Box 11.2 (continued)
Finally, there are statistical measurements that suggest how likely it is that the coefficient is truly different from zero and not just a random fluke. If that likelihood is high enough (if such a fluke would happen only about once out of twenty studies with similar data) the coefficient is called **statistically significant**.

This has been a technical detour, but a useful one, since a large percentage of economic research follows an approach along these lines. The real reason you are still reading this, however, is because the topic is sex, and you want to hear what they found. Now you will find out:

As you would expect, the coefficients for income and sex were both positive and significant. Once Blanchflower and Oswald had determined their size, they could ask, how much of a reduction in yearly income would it take in this equation to exactly offset a particular increase in the frequency of sex—say from an average of once a month to once a week? The answer was approximately $50,000: on average, people will report themselves as equally happy if they have either the extra sex or the extra money.

They also found other results of interest. Marriage, which increases the average frequency of sex (they asked) is worth $100,000, while divorce deducts $60,000. (Better to have loved and lost. . .) Holding the frequency of sex constant, having fewer sexual partners produces more happiness, and being gay (again holding all else equal) has no impact one way or the other. The researchers caution us, however, that they were unable to determine that it is sex that makes for happiness, rather than a pattern in which people who are already happy find sexual partners more easily.

For the full story, read "Money, Sex, and Happiness: An Empirical Study" by David G. Blanchflower and Andrew J. Oswald, National Bureau of Economic Research Working Paper No. W10499 (May, 2004).

Now on to an important consumption-related issue to test whether happiness has something to offer that utility doesn't. One of the pressing economic questions facing the United States is whether the spectacular increase in suburban and exurban development in recent years is desirable or not. (An exurb is an area beyond the suburban fringe populated primarily by residents who commute to the suburbs or cities.) More than half of all Americans now live in suburbs, displacing large tracts of what was once farmland. Highways leading into and out of the major cities are choked with traffic during rush hour, and the daily commute can take as much as two hours in each direction. Is this a problem?

If consumers are rational utility-maximizers, perhaps not. The main cost faced by someone who chooses to live far from where she works is the time and expense of commuting; the benefits are having a pleasant neighborhood with good schools, a desirable house and lot, etc. When choosing where to live and where to work, our

hypothetically rational individual will factor in all these considerations. She would not choose to live in a distant suburb, for instance, unless she calculated that the extra commuting burden would be at least made up by all the other advantages. Thus, if utility theory is our guide there is no case for public intervention—at least, not to rescue her from her rush hour misery. (There may be environmental or other considerations, of course.)

The utility maximization argument is based on the *assumption* that people maximize utility. Since there is no way to measure it directly, there is no way to test this assumption. On the other hand, happiness, at least as reported in surveys, is measurable. A recent study found that, even after taking into account all other factors, such as those the rational person would consider, longer daily commutes are associated with lower happiness scores. It appears as though, when choosing jobs, houses and apartments, people systematically underestimate how miserable their commutes will make them.

Does this mean that a government agency should take this decision out of the hands of private citizens? Not necessarily, but there are other possibilities. The happiness finding, if it holds up in other studies, might provide a justification for taxes, such as on gasoline, that favor people who live closer to their jobs. This would give people an extra incentive to make the choices that will, on average, increase their happiness anyway. At the very least, happiness research in this case neutralizes the argument that public policies on land use should not be adopted because they interfere with the choices people have made in a rational, fully informed way.

We will encounter happiness research again in the volume on macroeconomics, when we ask whether per capita Gross Domestic Product can serve as an indicator of general well-being. For now, its main purpose is to demonstrate that practical alternatives to utility theory exist and are being employed by economists. The field is still new, however, so it will probably not send quite the same set of messages several years from now when more results are in.

11.6 Capabilities

Happiness and utility are both essentially subjective concepts; they ask and try to find out about the feelings people have in their mind. It is also possible, however, to approach well-being from the more objective standpoint of evaluating the goods, resources and opportunities people actually have, whether or not they say they value or are even aware of them. The problem is to define what these valuable things might be in a way that is specific enough to be measurable, but also general enough to apply across individual and cultural differences.

Nobel laureate Amartya Sen has attempted to do this in his theory of **capabilities**, which he developed in conjunction with the philosopher Martha Nussbaum. Their idea can be traced back to Aristotle, who argued that, through observation of a number of communities, some of them successful and some not, it would be possible to determine objectively what conditions would have to be met

for human beings to "flourish". When Aristotle wrote this he had in mind a smattering of small Greek city-states; today we have the experience of the entire world to draw on. Is it possible to draw up a list that will work in Los Angeles, Hong Kong and Johannesburg?

The trick, according to Sen, is to put aside specific goods and concentrate on fundamental human functionings. To have the ability to carry out those functions is to have, in Sen's terms, the capabilities they require. An obvious choice would be nutrition: without specifying exactly what people should eat, it is clear that they need a sufficient nutritional intake to function without hunger or health impediments. Other basic needs enter in a similar way. But Sen goes further and argues that essential capabilities also include the social and cultural aspects of life. He points, for instance, to Adam Smith's observation that all people need the wherewithal to "appear in public without shame". This may mean one type of clothing in Bengal and another in Italy, but the same fundamental capability is at stake. Comparable arguments can be made about types of education, access to transportation and other resources: what is being measured, in principle, is not any particular set of goods and services but the capabilities of individuals to participate in various ways in the life of their communities. In fact, some of the capabilities on Sen's list are not economic in the conventional sense, but political and cultural, such as freedom of expression and the democratic accountability of government.

The capabilities approach was originally developed under the auspices of the World Institute for Development and Economic Research, WIDER, an affiliate of the United Nations, and during the past two decades an effort has been made by several international agencies to translate the theory into quantitative indices. Its greatest impact has been at the level of society-wide evaluation: what overall effect does a national policy strategy have on the capabilities of its citizens? How can we rank different countries according to their success in meeting economic and social goals? At this point it is not refined enough to apply to the narrower questions that are asked of particular industries and products. It does serve to remind us, however, that there is a case to be made for evaluations of well-being that take into account how people actually live, and not just their self-perceptions. Indirectly, it endorses the position of specialists in fields outside of economics, such as public health and education; the goals they pursue can be justified normatively on their own terms, now understood as capabilities, without being translated into the synthetic amalgams of utility and happiness.

The Main Points

1. The main purpose of demand theory is to support normative economic analysis; it proposes a relationship between market demand and the well-being of people who purchase goods and services. It has little to offer for positive analysis since it takes preferences as given: it doesn't examine why people have the preferences they have, or what factors might cause them to change their preferences.
2. Utility is proposed as the "substance" of well-being. It is something of a black box, being whatever it is that people hope to acquire by purchasing items for sale. It does **not** signify usefulness in particular, however.

3. The "law of diminishing marginal utility" states that, as one buys more and more units of a particular good, the additional utility acquired from one more unit declines. This is seen as the psychological basis for the "law of demand", which is that a lower price normally results in a higher quantity being demanded in the market.
4. The world of market demand—prices offered and quantities purchased—is visible; the world of utility, to the extent that it exists, is invisible. The "exchange rate" that converts the first into the second is the marginal utility of money. Individuals differ in their marginal utilities of money for two general reasons: the more money one has, the less additional utility one is likely to get from having a little bit more, and people differ in how much they value the things money can buy compared to the things it can't.
5. Consumer surplus is the difference between what individuals would be willing to pay for an item and what they actually have to pay—the market price. It is common to measure this in monetary terms, although, strictly speaking, consumer surplus in utility cannot be inferred from market demand because of differences in the marginal utility of money.
6. The analysis of market demand based on utility theory permits a formal statement of the conditions that must hold if one is to accept the Market Welfare Model interpretation that the demand curve represents the marginal benefits to society: (1) the utility theory of benefit is accepted as correct, (2) the demand curve is derived from willingness to pay, (3) willingness to pay is an accurate measure of marginal utility, (4) the marginal utility of money is equal across consumers, and (5) there are no impacts of consumption other than those measured by willingness to pay.
7. Research in behavioral economics casts doubt on the utility theory developed in this chapter. Findings include faulty self-perception on the part of consumers, the existence of multiple "mental accounts" that lead to inconsistent behavior, status quo bias, misperception of risk, and poor forecasts of the benefits derived from the goods people purchase.
8. Because of this there has been an upsurge of interest in an alternative measure of well-being, self-reported happiness or satisfaction. Evidence supports the notion that answers given by individuals to questions about well-being in surveys is consistent with objective indicators, like displays of emotion and neurological response. Research finds, however, that people do not maximize self-reported happiness corresponding to the way they are supposed to maximize utility.
9. An alternative approach to well-being is the theory of capabilities, which proposes that there are universal human needs and activities which economies can support to a greater or lesser extent.

▶ Terms to Define

Capabilities
Coefficient
Consumer surplus
Demand schedule
Law of demand
Marginal consumer
Marginal utility
Marginal utility of money
Paternalism
Statistical significance

Questions to Consider

1. Does a consistent liberal (in the sense we are using in this book) have to be opposed to laws banning drugs like marijuana, cocaine and ecstasy? Are the laws we currently have paternalistic? Do you get the same answers to these two questions if they asked about regulations taking ineffective medications off the market?
2. Does the concept of consumer surplus describe the benefits you get from what you buy? Think of a recent purchase: how would you compare the marginal utility of the good you bought with the utility value of the money you paid for it? Are there aspects of this purchase that don't fit comfortably with the consumer surplus model?
3. What do you think about the issue raised in Box 11.1? Should fines be set to equalize the monetary cost or the utility cost? In your answer, do the incentive effects of the fines play a significant role?
4. Discuss with a friend your relative marginal utilities of money. If it is greater for one of you, why?
5. A parcel of land adjoining a river is coveted by two groups. One consists of fishermen; they want the river to remain in a healthy condition so it can support fish, which they can then try to catch. The other is a mining operation that would dump tailings into the river, killing the fish. The market solution is to let them both bid on the land and have it sold to whoever expresses the highest willingness to pay. Can you extend this story in a way that brings differences in the average marginal utility of money between the two groups into the picture? In practice, how significant a factor is this issue likely to be in disputes between preservationists (like the fishermen) and developers (like the mining company)? Why?
6. Scan the list of consumer biases on pp. 229–239. How many apply to you? Have any of them had serious consequences?
7. Revisit the answers you gave to question 1. Do you think happiness research could have a role to play in these issues? In the first, the question is whether drug users are more or less happy because of their drug use; in the second, it's whether

pharmaceuticals that are deemed ineffective in laboratory experiments nevertheless contribute to greater happiness on the part of those who take them.

8. Sen says that, among the capabilities that all should have, an important one is access to education sufficient to permit everyone to participate effectively in political debate. What level of education is that in the US? How close are we to meeting that goal?

Appendix: Indifference Curves

The relationship between the price world and the utility world in this chapter is explained by juxtaposing two diagrams, Figs. 11.2 and 11.3. I appealed to your intuition to establish the logical connection between them. If you are still skeptical and want a more worked-out proof of the relationship, this appendix may provide it. It assumes somewhat more familiarity with analytical geometry than does the main body of the text.

The entire analysis is conducted at the level of a single individual. For simplicity, we will assume that only two goods, bread and cheese, are available for consumption, although nothing we say would have to be altered in a many-product world. Let's also assume that, for each possible combination of particular quantities of bread and cheese, the consumer is able to attach a utility value. Thus, five loaves of bread and two pounds of cheese have one value; three loaves of bread and four pounds of cheese have another, which may be less, more or equal.

Each of these combinations can be depicted as a point in plane defined by two (orthogonal) axes, one for bread, the other for cheese. Figure 11.6 on the following page locates two such points, (five loaves, two pounds) and (three loaves, three pounds).

Each point in Fig. 11.6 exists in two dimensions, bread and cheese. What we are interested in is utility, however, and that constitutes a third dimension. You could say that every point ought to transmit three pieces of information: the amount of bread, the amount of cheese and the amount of utility. From this perspective, utility constitutes a third dimension, rising up out Fig. 11.6; we show this in Fig. 11.7 as an optical illusion:

At point D the consumer has a bit more bread but a lot less cheese than at point E. The difference in utility is represented by showing E at a higher elevation on the vertical utility axis. If every possible combination of bread and cheese were given its corresponding utility value and plotted in three dimensions in Fig. 11.7, the result would look like the **utility surface** loosely pictured by the shaded area above. Any point along this surface would be traceable in three dimensions: bread, cheese and utility.

The utility surface can be compared to any surface in three dimensions. Consider a landscape, for instance. Suppose it encompasses a mountain rising out of a surrounding plain; this too is a surface in three dimensions, west to east, south to

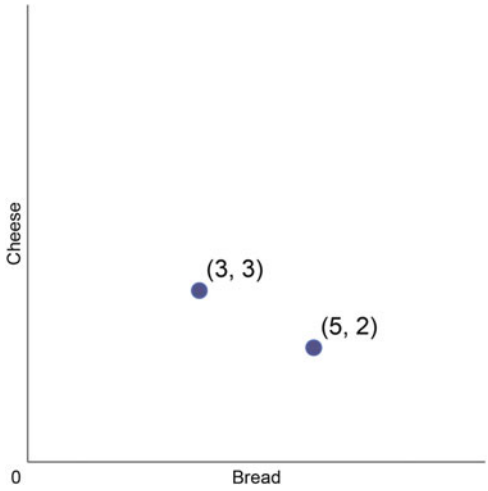

Fig. 11.6 Two combinations of bread and cheese. Loaves of bread are measured along the horizontal axis, pounds of cheese along the vertical axis. Two points are represented, with the quantity of bread given first in parentheses

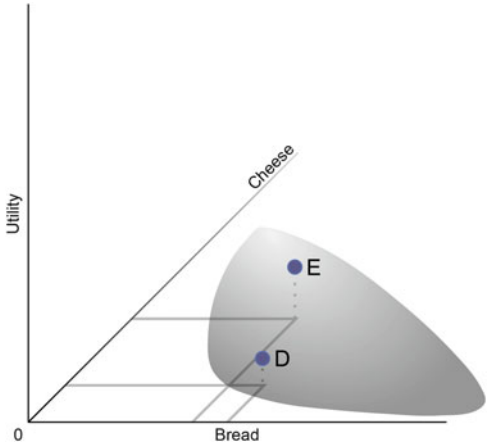

Fig. 11.7 Combinations of bread, cheese and utility. Two points, D and E, represent different combinations of bread and cheese. E is at a higher elevation on the utility axis than D, meaning that it provides more utility

north, and down to up. We could pick any two points and, if we knew their exact location and place on the mountain slope, we could say which one was higher.

Three-dimensional landscapes can be depicted in two dimensions with the aid of **contour lines**, which connect points of equal elevation. This gives us a contour or **topographical** map, familiar to hikers and other outdoors people. In Fig. 11.8 I have reproduced part of the contour map for Mt. Ranier, a large volcanic peak near my home in Washington State. (Yes, it's an active volcano.) The contour lines tell us how we would have to walk if we were following the exact contour of the slope, neither gaining nor losing elevation. By looking at how they are placed, we can determine which way the slope is facing, which way is up and even how steep it is (by the distance between contours). This is very useful for those traveling in this region.

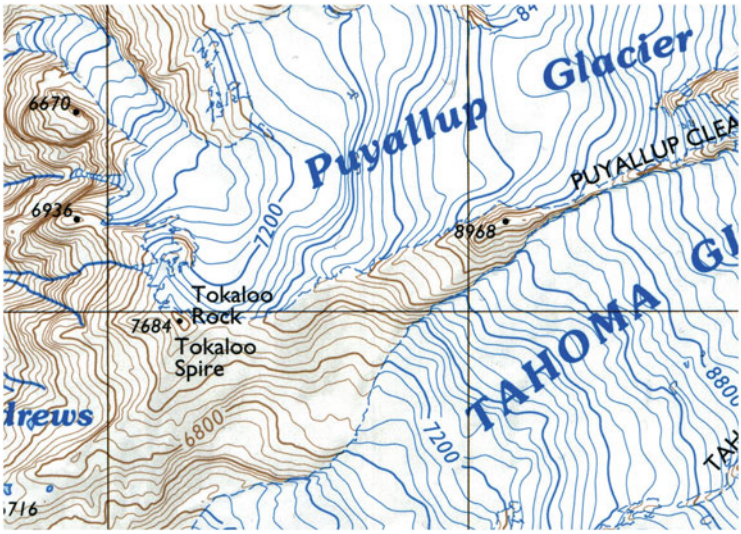

Fig. 11.8 A contour map representing a portion of Mt. Ranier

The consumer's utility surface can also be mapped in two dimensions by using **indifference curves**, lines connecting combinations of bread and cheese with equal levels of utility. The analogy is precise: indifference curves are to three-dimensional utility surfaces as contour lines are to three-dimensional mountain slopes. They are called indifference curves because the consumer is said to be indifferent between alternative combinations of goods along the same curve. Incidentally, representing utility surfaces in this way has an advantage beyond graphic convenience, since individuals could construct their own indifference curves simply by asking themselves, "Do I prefer the combination of goods in this point to the combination in that one, or vice versa, or am I indifferent?" Every time they decided they were indifferent, they would put those two points on the same indifference curve. Eventually, after asking this of every possible pair of points, they would be able to construct a complete **indifference map**, the utility equivalent of the outdoor lover's topographical map. This sounds like a lot of comparison (and it is), but it is less fanciful than supposing that people good assign actual utility numbers to each point. ("I like this combination of bread and cheese; I think I'll give it a 91.") In other words, using the language introduced earlier in this chapter, the indifference map can be constructed ordinally and not just cardinally.

Figure 11.9 shows a portion of such a map, picturing three different indifference curves for a consumer. Which do you suppose represents the lowest utility and which the highest? (You should stop and think about this for a moment.) The only way to know for sure is to introduce another assumption, that the individual always prefers more to less of every good. (The technical name for this is the *nonsatiation principle*.) As a universal statement it is not very appealing, and in fact puts us dangerously close to Imelda Marcos territory, but it is probably true for a majority

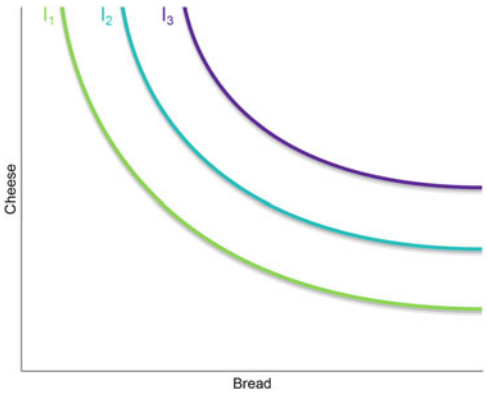

Fig. 11.9 Three indifference curves for an individual consumer acquiring bread and cheese. I_1, I_2 and I_3 are three indifference curves in order of lowest to highest utility. Each represents combinations of bread and cheese giving equal utility to some consumer

of goods the majority of the time (at least for those of us whose limited incomes regularly force us to buy less than we might otherwise). Once we accept this principle, however, we know that the indifference curves to the northeast must be at a higher utility level than those to the southwest. The reason is that, for every point on a lower curve, such as I_1, there is another point on another curve further from the origin, like I_2, that has at least as much of one good and more of the other, or indeed more of both. According to the nonsatiation assumption, the second point must be preferred to the first, and therefore all the points of equal utility to the second are preferred to those equal to the first.

If preferences are consistent, the indifference curves can't cross. If this were to happen, then one point would be on two different indifference curves, meaning that the combination of goods it represents would be equal to two other combinations which were not equal to each other. This would be a logical inconsistency—as it would be if two contour lines crossed on a topographical map.

Note that the indifference curves drawn in Fig. 11.9 are convex to the origin—why? To answer this, we need to understand more precisely what the slope of the curve signifies. The slope of an indifference curve at any point equals the slope of a straight line tangent to the curve at that point. This is depicted in Fig. 11.10 on the next page.

The tangent has the property in the vicinity of D that utility is not changed for small movements to or from D. This is because, as the movement along the line approaches D, it approximates movement on the indifference curve around D. (As you learn in the calculus, this is exactly true as the movement along the tangent from D approaches zero if the curve is smooth and continuous, which we assume it is.) Moreover, movements along the indifference curve do not alter utility, as we know from the definition of indifference. Put these two considerations together, and you can specify the equation for the tangent:

$$\mathrm{MU(B)}^{*}\Delta B + \mathrm{MU(C)}^{*}\Delta C = 0 \tag{11.2}$$

This equation reads "the marginal utility from bread times the change in the bread consumed, plus the marginal utility of cheese times the change in the cheese

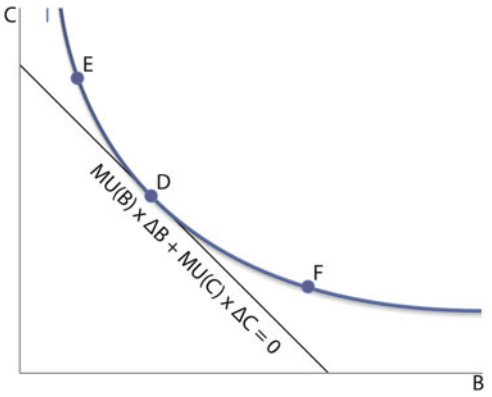

Fig. 11.10 The slope of an indifference curve taken at a single point. The slope of the indifference curve at point D is given by the slope of the line tangent to it, whose equation signifies that utility is unchanged. This slope would be different at points E and F

consumed, equal zero." Movement up and down this line increases the amount of one good at the expense of the other; this leaves utility unchanged only if the additional utility gained by increasing one exactly offsets the additional utility lost by decreasing the other, and marginal utility means simply the additional utility plus or minus with the change in a single unit (loaf of bread, pound of cheese).

The slope of this line is the change in its vertical component divided by the change in its horizontal component, or, in this case, ΔC/ΔB. We can calculate this from Eq. 11.2 with a little algebra:

$$-\mathrm{MU(B)}^{*}\Delta B = \mathrm{MU(C)}^{*}\Delta C \tag{11.3}$$

Divide both sides by MU(C) * ΔB:

$$-\mathrm{MU(B)/MU(C)} = \Delta C/\Delta B = \text{slope of indifference curve at D} \tag{11.4}$$

Let's explore the meaning of this result. The slope is negative, because it takes more of some good to make up for less of the other if utility is to remain constant. This negative amount is *the inverse ratio of marginal utilities*; that is, if ΔC/Δ B = x, then –MU(C)/MU(B) = 1/x. If, for instance, the marginal utility of bread is twice that of cheese, utility will remain constant if the reduction is bread is half the increase in the amount of cheese. This is not very profound, and it follows directly from Eq. 11.2, but what is interesting is that it enables us to infer the ratio of marginal utilities from the changing slope of the indifference curve, if we know what to look for.

Consider another point, E. As drawn, the slope of the indifference curve is steeper here than at D; it takes a bigger increase in the amount of cheese to make up for a smaller decrease in the amount of bread. From Eq. 11.4 this means that the ratio of the marginal utility of cheese to that of bread has gone down. Conversely, at point F it takes a smaller increase in cheese to make up for the loss of bread, and the marginal utility of cheese must therefore have risen relative to that of bread. Put it all together and what you see is that, as the consumer specializes in cheese relative

to bread (in the northwest portion of the diagram), the marginal utility of cheese relative to bread is falling, and similarly if the consumer specializes in bread relative to cheese (in the southeast portion). *The more the consumer specializes in the consumption of a particular good, the less relative marginal utility he or she gets from it*. This is nothing other than the law of diminishing marginal utility, applied to a situation in which two goods are being considered in relation to each other.

We can now answer the question we asked ourselves about a page and a half ago: the reason the indifference curve was drawn convex to the origin was to have this property of diminishing marginal utility. If I had drawn it concave to the origin (imagine the indifference curve in Fig. 11.9 attached to the tangent at D by a ring, and flip it over to the other side of the line), we would have depicted increasing marginal utility—the more I have, the more I want it—instead.

Now we will add one more wrinkle, a fixed amount of money and prices for the two goods. The consumer cannot buy an unlimited amount, but must now figure out how to apportion the money between bread and cheese in order to maximize utility. How much of each item to buy will depend on both factors—how much money is available to spend and how much each costs.

The limitation on how much the consumer can buy is called the **budget constraint**, and it is portrayed in Fig. 11.11 on the following page. If all the money is spent on bread, B_{max} is the amount of bread that can be bought; if all of it is spent on cheese the amount is C_{max}. Intermediate amounts can be purchased by buying more of one and less of the other; these are on the straight line connecting B_{max} and C_{max} because the tradeoff (and therefore the slope) is unchanged due to the prices remaining unchanged. The consumer can buy any combination of bread and cheese, provided it is to the southwest of the line or just on it. Anything to the northeast is unaffordable. The equation for this line simply says that the consumer spends all available money:

$$B^*P_B + C^*P_C = Y \tag{11.5}$$

where B and C are the amounts of bread and cheese purchased, P_B and P_C are their prices and Y is the amount of money to be spent.

Visualizing the best option for the consumer will be easier if we return to the example of a topographical map. Suppose some of the land in Fig. 11.8 is on private land and some in the Mt. Ranier National Park, with the dividing line as shown in Fig. 11.12; what is the highest point on private land in the region depicted in the map? The answer is, where the boundary line just barely grazes a contour line, in this case at 7,200 ft: that will be the highest contour attainable if you have to stay out of the park. This is given by point A in the map.

The same logic holds for our indifference map. If we superimpose our budget constraint on the original indifference map in Fig. 11.9, we arrive at Fig. 11.13:

The highest indifference curve that can be attained is I_2. Points along I_3 would be preferred, but they are out of reach. The budget constraint barely permits the consumer to select point A, representing B* loaves of bread and C* pounds of cheese.

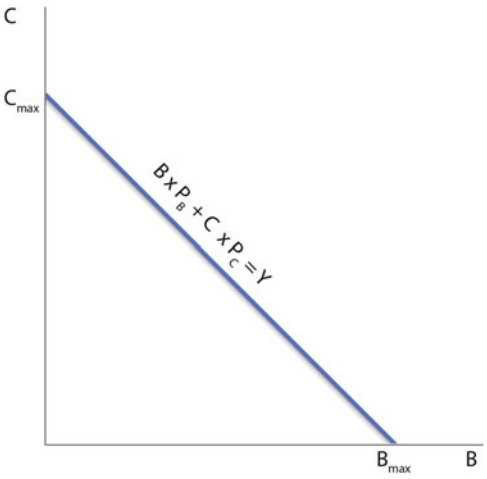

Fig. 11.11 A budget constraint for bread and cheese. Given an amount of money Y and prices P_B and P_C for bread and cheese respectively, the consumer cannot purchase combinations of bread and cheese beyond the budget constraint drawn above

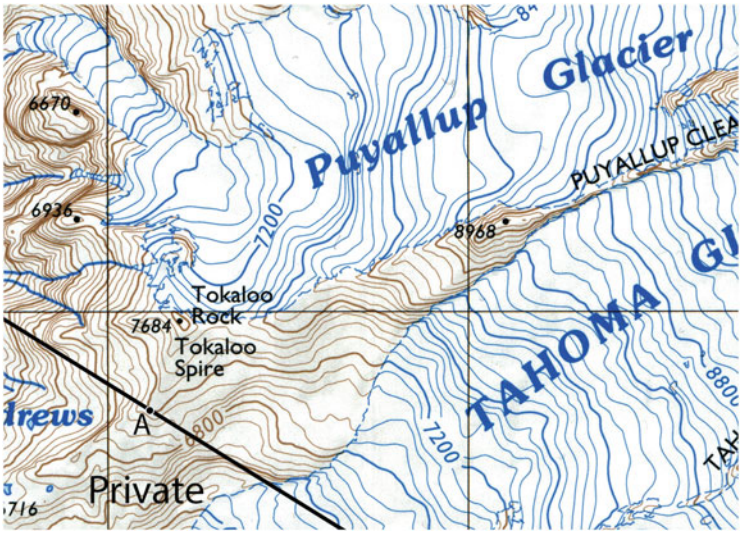

Fig. 11.12 Mt. Ranier topographical map with hypothetical line separating private from public land. Land to the SW of the boundary is private land; land to the NE is public. The highest point that can be reached without entering the park is *A*, where the boundary is tangent to the contour line

Of course, we have already worked out what it means for a straight line to be tangent to a point along the indifference curve; its slope tells us the inverse of the ratio of marginal utilities. Now, however, we have a new piece to add, because the budget constraint given in Eq. 11.6 must obey the condition

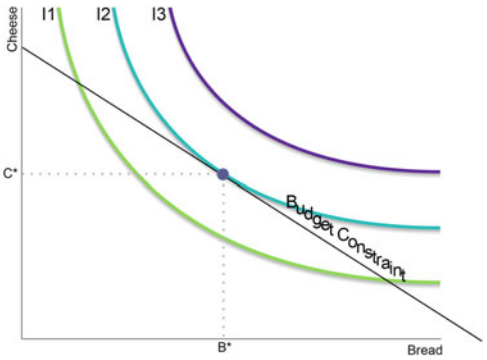

Fig. 11.13 Three indifference curves for an individual consumer acquiring bread and cheese, plus a Budget Constraint. When the consumer faces a budget constraint as above, the highest indifference curve that can be reached is I_2, where the budget constraint is tangent at point A. The consumer will buy B* loaves of bread and C* pounds of cheese

$$\Delta B^* \mathrm{P_B} + \Delta C^* \mathrm{P_C} = 0 \tag{11.6}$$

That is, if you increase the amount of bread purchased, that will deduct $\Delta B * P_B$ from your budget, which can only be made up by saving $\Delta C * P_C$ on less purchases of cheese. But, as before, $\Delta C/\Delta B$ is the slope of the line, so we need to rearrange Eq. 11.6:

$$-\Delta B^* \mathrm{P_B} = \Delta C^* \mathrm{P_C} \tag{11.7}$$

Now divide both sides by $\Delta B * P_C$:

$$-\mathrm{P_B}/\mathrm{P_C} = \Delta C/\Delta B = \text{slope of budget constraint} \tag{11.8}$$

The slope of the budget constraint will be negative (the line is downward-sloping), so it will equal the inverse ratio of the prices. The steeper the slope, the greater the shift in cheese needed to offset a shift in bread, and the lower the price of cheese relative to bread.

Since the slope of the budget constraint is also the slope of the indifference curve at A, we can bring Eqs. 11.4 and 11.8 together:

$$\begin{aligned} -\mathrm{MU(B)}/\mathrm{MU(C)} &= -\mathrm{P_B}/\mathrm{P_C} = \Delta C/\Delta B = \text{slope of budget constraint} \\ &= \text{slope of indifference curve} \end{aligned} \tag{11.9}$$

The ratio of marginal utilities will be equal to the ratio of prices when the consumer maximizes utility subject to a budget constraint. Logically, we can imagine the prices being imposed from the outside (from the market) and the consumer adjusting purchases so that the equality in Eq. 11.9 is established. For instance if the price of cheese rises relative to that of bread, the slope of the budget constraint will become flatter. This will lead to a new optimum purchase

combination somewhere to the right of point A, presumably on a different indifference curve. Since the new purchase decision will involve increasing the amount of bread (which is what it means to be to the right of A), the marginal utility of bread will diminish relative to that of cheese. This process is complete when those two ratios (prices and marginal utilities) are once again equal. Incidentally, notice that the law of demand works in this model: increasing the price of one good leads to a shift in purchases to the other.

The equality of the price and marginal utility ratios is exactly the result in the two-good case that corresponds to the relationship between marginal utility and willingness to pay in the one-good case. Note that Eq. 11.9 does not say that the marginal utility from either good equals its price. This would be meaningless, since the two are measured in entirely different units. It does say that, whatever the ratio between the price and marginal utility of a good, that same ratio applies to the other good. This is because, by dividing both sides of Eq. 11.9 by $-P_B/MU(C)$ we get

$$MU(B)/P_B = MU(C)/P_C \tag{11.10}$$

Consider this: we have developed the model using two goods, bread and cheese, but we could have done it for any two goods, or any larger number of goods, for that matter (although we would not be able to use two-dimensional diagrams for more than two goods). What if one of the goods were money itself? Then the price of money would, of course, be exactly one. (It costs exactly one dollar to buy a dollar if you're paying attention.) This means that the left side of Eq. 11.10 becomes simply the marginal utility of money. Since this equals the ratio of the marginal utility of some, or any, other good to its price, we are directly back in the world we created in the first half of this chapter.

Equation 11.10 can be given a somewhat different interpretation as well. You can read it as saying that, if you divide the marginal utility you get from any good by the price you pay for it, the result will equal the same operation for any other good. More bluntly, the marginal utility you get per dollar will be equalized across all goods when you maximize your utility while remaining within your budget constraint. There is an intuitive logic to this. If it were not the case, you could take a dollar out of a good whose marginal utility per dollar was low and reallocate it to something else where it was higher. By doing this, however, you are increasing your purchase of the higher marginal utility good which means, according to the law of diminishing marginal utility, that it will go down. The process continues until it has dropped to the level of the good you are buying less of—which of course, because you are buying less of it, will give you greater marginal utility. This notion is a bit idealized, because it assumes that you can make dollar-by-dollar reallocations, whereas real-world goods often have to be purchased in lump sums of many dollars. With "lumpy" purchases of this sort, you would never attain the perfection of Eq. 11.10, but you would come as close as you could.

The interesting thing about Eq. 11.10 is that it ought to hold for *every* individual in the market, since all face the same set of prices. Thus we come to the same point as in this chapter: if we could just say that the marginal utility of money were equal for all individuals, or even if we knew the ratio of each individual's marginal utility of money to the average, we could derive utility information for the entire population just by observing market prices. It is difficult to justify this step, however, so we are pushing the outer limits of what the theory has to say to us beyond the level of a single individual.