CHAPTER 2

Individual Choices: Explaining Food Consumption and Production

2.1 Consumer Choices: Food Preferences and Dietary Intake

2.1.1 Motivation and Guiding Questions

People choose what to do from a limited set of options. What determines those options, and how does each person decide which of them to choose? Why do people at the same place and time often eat similar foods, while others have very different dietary patterns? And most importantly, to guide intervention, what can an outside observer infer from observed choices about a person or population's level of wellbeing, in a way that might guide intervention to improve outcomes?

In the health sciences, researchers and practitioners often answer these questions using psychology and a social-ecological approach to health behavior. Nutritionists and dietitians draw on the health sciences to explain food choice as the result of each person's individual response to their circumstances, based on the individual's biological needs, psychological needs or social condition in the context of their household, community and broader environment. Nutritional epidemiologists often refer to a person being 'exposed' to certain foods, in the same way that they might be exposed to other factors influencing their health such as viruses or air pollution.

The health behavior approach can be very helpful in clinical practice or other settings, but it is focused on providing guidance towards healthier choices. The economics approach to food choice aims to explain and predict observed food choices, whatever they may be, in a way that allows us to infer something about the population's preferences. Both health behavior and economics research start with the dignity and agency of each individual,

W. A. Masters and A. B. Finaret, *Food Economics*,
Palgrave Studies in Agricultural Economics and Food Policy,
https://doi.org/10.1007/978-3-031-53840-7_2

recognizing that every person responds to their circumstances in their own unique way. Both then observe that human biology and other factors introduce enough commonality that whole populations often behave somewhat similarly in response to different circumstances.

In economic models of consumer behavior, the underlying structure behind food choice is the idea that people have selected what we observe from a limited set of available options, in pursuit of their individual goals. In economic terms, goals are represented as preferences. These preferences are sometimes described as a population's *utility function*, meaning the usefulness of each thing in pursuit of the population's various goals and aspirations. A person's preferences describe how, in terms of Alfred Marshall's original definition of economics mentioned in Chapter 1, a person uses 'material requisites' to form their 'wellbeing'.

Some things may be consumed for their own sake, while others may be instrumental for some other purpose such as future health. The options from which a person can choose are constraints on their wellbeing. For food choice, economists illustrate those options in terms of relative prices (meaning the cost of choosing one thing instead of other things) and total income (meaning the sum of all things that a person could afford to choose). Health behavior interventions generally aim to alter preferences, while economic interventions often target prices or income.

In the graphical approach to consumer choice, each person's preferences are shown as indifference curves, where higher levels of those curves represent a more preferred outcome. The person's constraints are shown as budget lines, where higher levels of that budget line represent a larger total income or potential level of expenditure, while the slope of that budget line shows the relative price or cost of each unit along the X axis in terms of the number of units required along the Y axis. That 'rise over run' of the budget line is constant, whereas the slope of the corresponding indifference curve can vary. This section presents a unified economic framework for understanding food consumption decisions, to analyze how preferences shape food consumption when prices or incomes change and explore the evidence on what people actually eat around the world in response to differences in preferences, prices and income.

Our eating decisions are among the most frequent choices we all make. Most people eat multiple times per day, under different circumstances over time. The resulting dietary patterns are a major determinant of cardiometabolic disorders including diabetes and hypertension as well as several types of cancer. The severity of infectious diseases is also affected by dietary patterns, as poor nutritional status can limit immune response and worsen outcomes from all kinds of illness. Children are affected by their parents' diets, not only during pregnancy but throughout life, and poor dietary quality at any age can have personal, societal and intergenerational health consequences.

Every person has their own unique food preferences, with strong links to our psychological and moral or cultural wellbeing. Some food preferences depend on the biology of taste and texture, but people may also seek out food that is thought to be healthier for us and others, and contribute to other goals involving climate change and the environment, or community and social justice. Readers of this book will include people who follow many different special diets such as vegetarians or vegans that are chosen for reasons involving health, sustainability and social justice, while others will follow low-fat diets that focus on protein and carbohydrates, paleo diets that limit carbohydrates or diets that avoid specific compounds such as gluten-free and lactose-free diets. Each of those dietary practices can be represented in the economics framework as an aspect of the person's preferences guiding their day-to-day choices among all the options they might otherwise have chosen.

In this section, we will examine how to explain diets as peoples' choices from among their options, and thereby investigate why food choices might differ between individuals. Even when people face similar food prices at their local grocery outlets they will choose different items, in part due to different levels of total income, but also due to different preferences at a given level of income and prices. Explaining and predicting those choices is possible only to the extent that preferences are stable to some degree, over time for the same person and among people in the same population. Economists aim to observe a sufficient range of choices under diverse conditions for whatever set of preferences is revealed. For example, if a population consistently chooses to eat an average of 5% more avocadoes when the price of avocadoes falls by 10%, that information would be used to characterize the *revealed preferences* of that population.

All observations are subject to measurement error, and even if choices and circumstances were perfectly measured, we would expect some unexplained variation in any set of choices. But when enough high-quality data are available, populations often reveal consistent preferences that allow economists to make predictions about their average response to changes in income or prices. For example, if a population with options A and B typically choose A, and when they have options B and C they typically choose B, economists predict that they would typically prefer A over C if given that choice. Consistency in this sense has been observed in a very wide range of settings. People do sometimes behave inconsistently by choosing C over A, but that would be the part of behavior that cannot be explained by past choices using revealed preferences.

The purpose of explaining behavior in terms of revealed preferences is not just for predictions about what people will choose when they have different options, but also to permit a kind of inference from those choices about the population's level of wellbeing. In the example above, circumstances that remove option A can be inferred to have reduced the population's wellbeing, in the sense of their own revealed preference for A over B or C. The population's own preferences may not be what other people would want for them.

For example, young children might choose to drink soda every day instead of water or juice, while their parents might know that the child would later regret that. In such cases, observers can see that the child's long-term best interests are best served by having parents who restrict their beverage options. Even adults might make food choices that do not reflect their own interests, if only because consumers cannot see and are sometimes misled about the healthiness of different options.

Revealed preferences serve a population's own long-term wellbeing only to the degree that people have experienced the impact of each option on their lives and choose among their options in a way that serves their lifetime goals. Since the impact of food choices on future outcomes may be unknown or misleading, food policies often prohibit false claims and require labeling to disclose what's inside each food. Labeling and education may not be sufficient to align choices with lifelong interests, so populations may prefer to have some ingredients or types of food be banned entirely. In any case each person's observed choices reveal something about how each thing serves their wellbeing, as described in this chapter.

In the section below, we will see how any set of consistent preferences can be described as having pursued the individual's highest available level of subjective wellbeing from their own perspective. In that sense, people can be said to have chosen the best or least bad of their options, based on what they have experienced or know about the consequences of each option. In other words, people make choices that are 'optimal' for them, 'maximizing' the utility or usefulness of their available resources in pursuit of wellbeing. This terminology is one of the several cases where economics differs from everyday language. In normal life, an 'optimal' outcome is the best it could possibly be, whereas in economics it is just the best of the available options for that person. None of the options may be good, so the optimal choice we expect to observe is the least bad of each person's options. And economists expect those choices to reflect all the person's goals, whereas everyday language might focus on just one goal. For example, a most medical professionals might think of an 'optimal' diet as maximizing health, whereas an economist would use the term to mean a diet that best achieves all the person's goals including health but also convenience and other aspirations.

By the end of this section, you will be able to:

1. Describe the economic determinants of food consumption choices;
2. Sketch indifference curves and budget lines to explain choices as points on a diagram;
3. Use the analytical diagrams to explain and predict change in food choices in response to change in prices, incomes and preferences; and
4. Describe strengths and limitations of the economics approach to explaining food choice.

2.1.2 Analytical Tools

The toolkit of economics is a set of mathematical models that we can build using lines and curves on a two-dimensional diagram. Each line or curve shows a relationship between two things, drawn with a shape and position that represents an equation between the two variables shown in the graph, holding constant all other variables.

A Model of Consumer Choices

The shape and position of each line or curve represents a set of facts about the world. For example, we will start with diagrams about an individual person's food choices in which preferences are shown with curves that always slope down and are bowed in, like the bottom-left corner of a circle O, or the bottom half of an opening parenthesis. The set of all such curves parallel to each other forms a nest like (((. We then draw the options among which they choose using a downward sloping straight line, whose position represents the person's income, and the slope represents the price they pay to consume one more unit of the variable shown on the horizontal axis. When different people shop at the same grocery store and face the same prices, their incomes are shown as parallel lines like \\\. The points where a curve just touches a line is a possible choice, and we use that system of simultaneous equations to explain observed choices, and predict the outcome of changing incomes, prices and preferences.

Notation and Specification of Variables on Each Axis

In this section we start our formal analysis by defining *goods* as anything for which more is better and less is worse. Most foods are goods in that sense, meaning that each additional unit adds something to the consumer's wellbeing. As we will see, increasing quantities are eventually subject to diminishing returns, and too much of a good thing can be bad, but the quantities consumed that we observe in practice are usually within a range over which additional (or 'marginal') units are desired in some way. Our analytical diagrams refer to the use of goods not because more is always better, but because people incur costs to obtain things, and those costs imply that people usually stop buying something when additional units are no longer desirable. Exceptions to that rule, when some people consume too much of a good thing, turn out to be an important aspect of food choice. That is one of many reasons why it is helpful to have a specialized textbook in food economics.

In this textbook, we begin building the toolkit of economics by representing individual behavior using the kind of diagram shown in this section. And in diagrams throughout this textbook, a solid black dot near the center represents the observed combination of things actually observed, while variables such as Q_x servings of product X are charted along the horizontal axis, and Q_a quantity of another things are charted along the vertical axis. Our goal is to explain why that quantity was chosen, predict what other choices might have been observed under other circumstances or a policy change, and evaluate whether such a change would improve or worsen this person's subjective well-being given their individual needs. Each food choice is made from a limited set of options shown by an area, line or curve, and changes in circumstances or policies shown by shifts in a line or curve lead to movements along another line or curve to a new food choice or other outcome.

The diagrams in this section of the book refer to quantities consumed by an individual person and have the observed quantities near its center because that gives us plenty of space along the axes with which to consider what other options might have been observed, under other circumstances. To show these comparisons visually must flatten the world into just two dimensions, so analysis using these diagrams begins by defining what is shown on each axis. For example, food economists and nutritionists are often interested in the total quantity of vegetables consumed along the horizontal axis, in contrast to other things along the vertical axis.

Indifference Curves for Consumption of Each Good

Analysis of food choice begins with the concept of an *indifference curve*, aiming to explain and predict consumption of something whose quantity is shown along the horizontal axis. Quantities of a food such as vegetables might be measured in servings (one tomato, two carrots or half an onion might all be considered one serving of a vegetable) or units of weight (such as ounces or grams) or volume (cups or liters). Nutritionists in the U.S. often measure fruits and vegetables in cup-equivalents, a hybrid unit that aims to capture just the solid dry matter in each food, while any kind of food can also be measured in terms of total dietary energy (in calories or joules) or grams of each macronutrient (carbohydrates, protein and fats). Quantities of something else along the vertical axis could refer to a particular thing, such as the quantity of fruit, and could be counted using the same units of measure as vegetables along the X axis.

For the diagrams in this section of the book, it is helpful for the vertical Y axis to add up the quantity of *all other goods and services* that a person might consume. The reason for this will be clear later when we consider how much of what's on the X axis a person can afford to obtain, which will be shown using the person's total income and the price of what's on the X axis relative to the prices of all other things. Adding the quantities of disparate things such as groceries and school supplies, restaurant meals and concert tickets cannot be done with physical units like cups or kilograms, but it can be done in terms of their monetary value. For that reason, one can think of all other things along the vertical axis as a stack of money, where more represents a larger quantity of all other things that could be obtained as shown in Fig. 2.1.

The purpose of Fig. 2.1 is to show all options that a person might find as desirable as the observed point labeled **O**, where they consume Q_x and Q_a. This observation might have come from a household consumption survey or dietary recall, in which the person reported having that many servings of

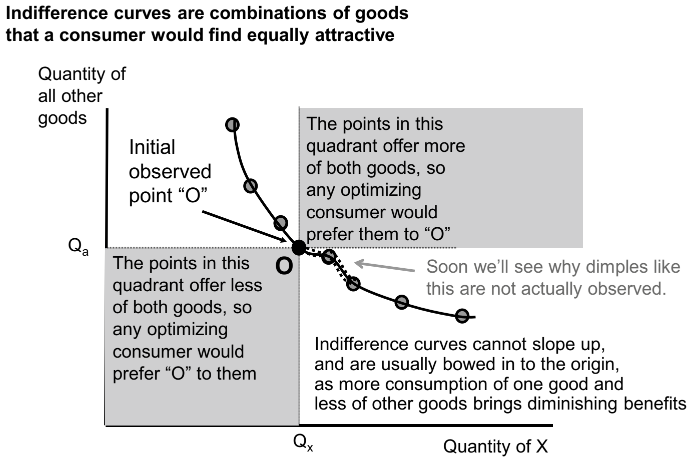


Fig. 2.1 Definition of the indifference curve

vegetables and that amount of total spending on all other things. To explain why the person chose this instead of some other possible combination of things, we must draw all possible alternatives to the observed point that could have been chosen instead.

Figure 2.1 shows how economists draw the foundations of food choice, using a curve to illustrate general principles about each person's needs and wants. The diagram shows all possible quantities that would provide this person with the same level of subjective wellbeing, using different combinations of Q_x (e.g., servings of vegetables) and Q_a (spending on other things). Each set of equally attractive options is called an *indifference curve* (IC), because the person whose preferences are shown in this diagram would be indifferent between all the points along that curve. The curve's specific location and shape will differ, but all indifference curves used in economics have two fundamental attributes:

First, indifference curves *always slope down* from left to right, to show that person would generally require additional quantities of the X good to compensate for less of all other things, if they are to maintain the same level of subjective wellbeing. This holds true as long as X is a good for which more is better. At extreme levels of X or Y, the curve might conceivably slope up, but we would not observe consumption choices in that region if X and Y are costly to obtain. Indifference curves that draw observed preferences will not slope up, but food economists understand that people may not choose quantities that are in their own long-run interests. Later we will compare the indifference curves that are revealed by a person's present choices with the preferences of their future self who might regret what was chosen. We will also discuss the consequences of being exposed to things that people themselves have not chosen, such as air pollution or contaminants in food, but the diagrams are designed to illustrate quantities of things that people have chosen to obtain.

Second, indifference curves typically slope down *with a decreasing slope*. The line becomes flatter with increasing quantities of what's on the X axis, reflecting how each additional unit of X is less valuable for this person's subjective wellbeing. That kind of decreasing returns in consumption gives indifference curves a bowed-in shape that mathematicians would call convex. As shown in Fig. 2.1, indifference curves may have regions that are not bowed in. The curve may have a bowed-out dimple where consuming a small quantity drives desire for more, so people are observed to consume either small or zero quantities to the left of the bowed-out segment, or large quantities to the right of the bowed-out segment. That idea was captured by a famous advertisement for potato chips, saying people 'can't eat just one', because eating one is likely to lead to eating more until some limit is reached.

Another example is how learning to cook at home builds skill that offers increasing returns up to a point, as practicing a few times makes future meals even better. At some quantity any person's subjective wellbeing from each additional unit will decline, resulting in a flatter indifference curve as quantities increase. Once people have experimented, their usual diets are such that additional quantities would yield diminishing returns, leading to a bowed-in shape for the indifference curves we draw around each point actually observed or predicted.

The downward sloping, bowed-in shape of each indifference curves follows from the fact that, around each observed point, the shaded region above and to the right of the observed point would have more of both things, so would have already been chosen if that had been possible, while the shaded region below and to the left of the observed point would have less of both things, so would not have been chosen instead of the observed point. Redrawing such quadrants around any potentially observed point reveals why the whole curve must slope down, as people's trial-and-error experiences with each food lead to the preferences we observe.

Having drawn one indifference curve through the observed point, we can see how other outcomes would provide different levels of wellbeing, as illustrated in Fig. 2.2.

Figure 2.2 has many indifference curves, each one representing different combinations of Q_x and Q_a that a person would find equally desirable. Higher levels of wellbeing are shown by points along a higher indifference curve, on which there might be more of everything that this person desires. Figure 2.2 shows how each level of wellbeing is illustrated by a curve that never crosses a lower or higher indifference level, unless the person has changed their mind to a different set of preferences as shown by the dashed curve. Along the solid curve all circles are equally attractive, but if this person's preferences change they might decide that the hollow triangle is as good as the solid dot, instead of the hollow circles which were their previous preferences.

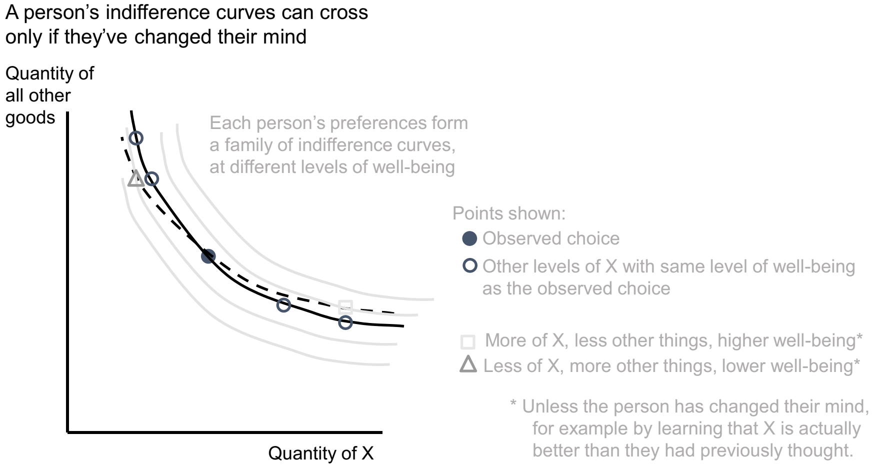


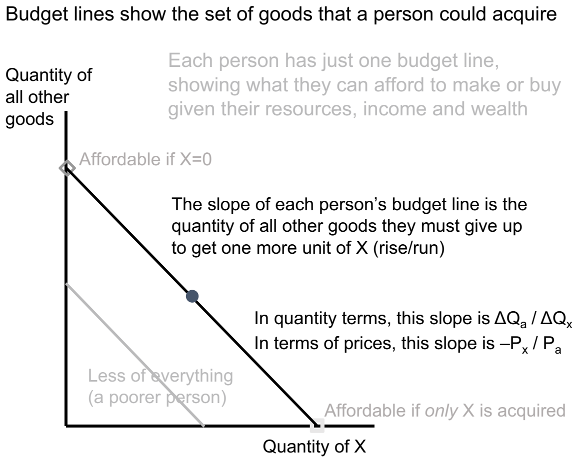
Fig. 2.2 Each person has many possible indifference curves

The purpose of economic models like Fig. 2.2 is to capture the predictable aspects of behavior. Having a stable set of preferences requires that a person's indifference curves not cross each other, so that each successive level of wellbeing is unambiguously higher or lower. If indifference curves were to cross, the person's preferences would lead to seemingly random switching for example from a circle to the triangle. In reality we observe some random behavior, for example when a person wants unexpected variety, but then we would draw quantities along the X axis as the fraction of time they want that thing.

A person's set of indifference curves can be imagined as topographic lines showing altitude on a map, or the lines of constant temperature on a weather map. The curvature of each line is important because it shows how rapidly the person's level of wellbeing changes as they increase consumption of each product. A gently curved indifference level implies that about the same quantity of all other things could substitute for the item of interest along the X axis, while a sharply curved indifference level leads to a narrower range of observed consumption. In extreme cases a person might have an L-shaped indifference level, implying that a fixed quantity of what's on the X axis is needed for each level of wellbeing, and any deviation from that leads to a different level of wellbeing. The meaning and use of indifference curves become intuitive as you practice sketching them, for example conducting imaginary thought experiments about your own food preferences.

Having established that a person's needs and preferences can be drawn as successively higher indifference curves, what level of wellbeing can a person reach? To answer, we need a different kind of line that shows the options from which they choose. Such a line illustrates all the possibilities that this person could afford, based on the money and time or other resources available to

Fig. 2.3 Definition of the budget line



them. The set of all options that a person could afford is known as a budget line, showing their possible total expenditure as drawn in Fig. 2.3.

The budget lines shown in Fig. 2.3 are drawn on the same axes used for indifference curves, but now the lines show all options that are equally affordable whereas indifference curves show the combinations that would be equally desirable. The difference is that budget lines have a constant slope. The slope of any line or curve is its rise over run, in this case denoted $\Delta Q_a/\Delta Q_X$ where Δ (delta) means difference from one point to the next, or change in Q_a for each unit of change in Q_X. In other words, the slope of a budget line is the quantity of all other things that must be given up to obtain one more unit of the thing along the X axis. If we imagine the quantity of all other things to be represented by a stack of money, then that slope is simply the price of X. We can also use 'price' metaphorically to mean everything that must be given up to obtain the thing of interest. Or, if the things on each Y axis also had their own price, we would need to divide the price of X by the price of Y to obtain the *relative price* of X. For that reason, the budget line's slope is generally written as $\frac{-P_x}{P_a}$. A negative sign appears before price because that is the amount of other things that must be given up to get a larger quantity of X, and a steeper budget line implies a higher cost of X. The slope of each budget line represents prices paid, while the level of each line represents the person's total income or expenditure. A budget line that is closer to zero shows how that person has fewer options, due to lower income so they can afford less of each thing. The vertical intercept of each budget line shows the person's income before buying any of the item along the X axis, and the budget line's horizontal intercept shows the quantity of X they could buy if they spent all of their resources on that item.

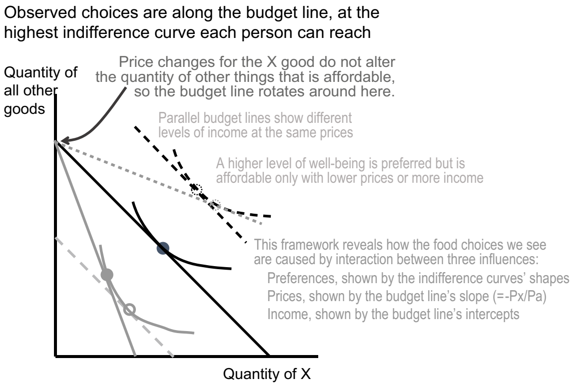


Fig. 2.4 What we observe is each person's preferred choice from the options they can afford

Having defined how budget lines show the options that are available and affordable for each person, and indifference curves show that person's preferences, we can now put together a complete model to explain what we observe and predict how changes in each causal factor would alter food choice. In the causal framework used by economists, each potential observation results from the individual having experienced different options and chosen what they prefer from their set of affordable options as shown in Fig. 2.4.

Figure 2.4 shows four possible points that differ from the observed solid dot in the middle, revealing how a higher level of wellbeing along the dashed indifference curve could have been reached with a lower price of X or a higher level of income, and similarly a higher price of X or a lower income could lead to lower wellbeing as shown by the lower indifference curve.

The general principle underlying each point we might observe is that the person's choices are based on their own experiences and knowledge of how each option might affect their wellbeing. Economists might say that the person has already optimized, choosing the best (or least bad) of their options, based on their own preferences. This way of explaining behavior is based on recognizing the limited agency of each individual, as they respond to the socioecological conditions around them. A change in the price paid for each X good, shown in Fig. 2.4 as rotation of the budget line around its Y intercept, would be the result of community factors such as the food environment, while a change in the person's level of income is generally a household characteristic.

In Fig. 2.4 the slope and curvature of the indifference curves have stayed the same for all four alternatives to the observed point, illustrating a situation in which the person's preferences have not changed. Later we will see how advertising, behavior-change programs and other interventions might alter preferences. Before that, it is important to note that most foods are not actually consumed at all, and when affordability or preferences change people switch from zero to significant quantities as part of an overall dietary pattern. That aspect of food choice and preferences is illustrated using Fig. 2.5.

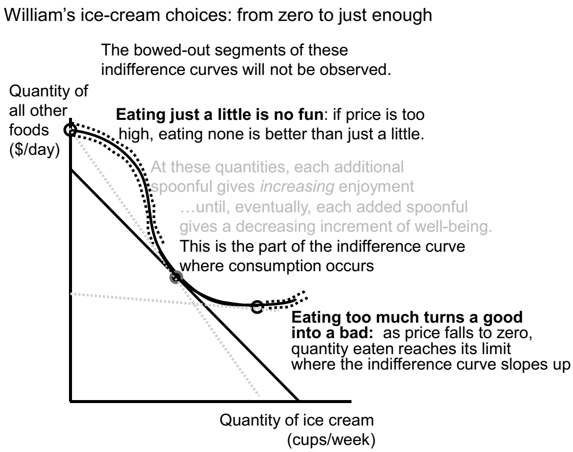


Fig. 2.5 What we observe is along a bowed-in portion of each indifference curve

Figure 2.5 has a large bowed-out segment on the left, whereas our initial indifference curve in Fig. 2.1 had a small dimple in the middle of the diagram. Both are possible. Having seen that people choose along their budget lines the combination that gives them the highest level of indifference, we can appreciate why bowed-out segments are not observed, and people often jump from zero or lower to higher quantities along a bowed-in segment of their indifference curve. The reason is that observers see only the outcome of each person's choices. By the time consumption is measured, the person has already experienced or imagined different options and chosen the best of what they can afford.

The example in Fig. 2.5 relates to William's high school job scooping ice cream. Now, as an adult, if ice cream were very expensive he would probably not eat any at all, because eating just a little makes the next bite all the more satisfying as shown by the steeper slope of the indifference curve when moving from zero to the right. William's experience with ice cream includes a time when it was basically free, so the price line was very flat but there was still a limit on how much he consumed. In other words, William's consumption of ice cream is always observed along the bowed-in and downward sloping of his indifference curves, precisely because he has experience with other quantities that led to the choices he now makes.

So far we have discussed consumption choices for an individual person. To clarify the story, it is helpful to imagine using one diagram to explain the different choices of multiple shoppers in the same supermarket as shown in Fig. 2.6.

The diagram in Fig. 2.6 shows the choices of nine different people, all of whom face the same market prices shown by the slope of their budget lines. Each of the shoppers has their unique level of wellbeing shown by their own indifference curve. The quantity of X that we observe shoppers having purchased is interpreted as having been chosen by them because it was the

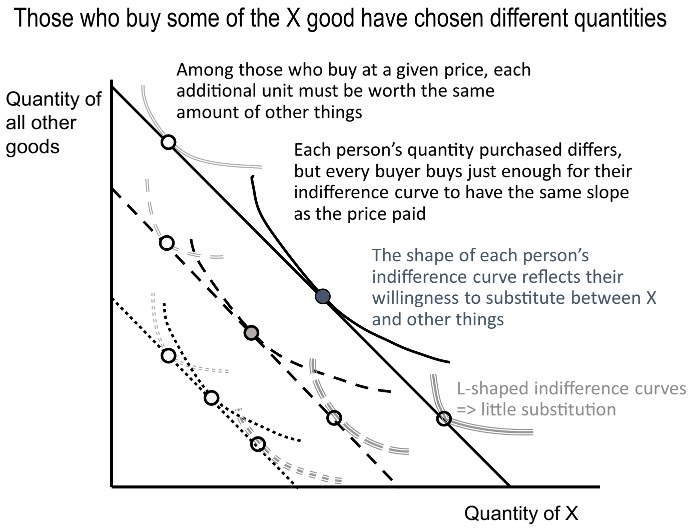


Fig. 2.6 People differ in their preferences and incomes, but face similar prices

best (or least bad) of their options on that day. For that reason, each indifference curve touches the person's budget line just once, because it shows the highest level of wellbeing they can reach. For simplicity we show only three levels of income, so all other variations are due to preferences. On the left of the diagram we see higher incomes corresponding to more purchase of other things, but no change in consumption of X, and in the middle we see higher income corresponding to more purchase of both things. All these outcomes are possible, with the economics framework allowing us to distinguish between income and preferences as a cause of the variation we observe.

An important observation from Fig. 2.6 is that each person has a different indifference curve, but at the observed quantities purchased all of those curves have the same slope. The reason is that people have moved along their budget line to their highest available indifference level, which is known in mathematical terms as a point of tangency between the person's budget line and their highest attainable indifference curve. As with any line or curve, the indifference level's slope is always its rise over run, which in this case would be written mathematically as $\Delta Q_a/\Delta Q_X$. At the highest attainable level of indifference, that slope is exactly equal to the relative price of X. If the indifference curve were steeper or flatter than the product's market price, the person would go back and adjust their purchase to where each additional unit purchased has a *marginal rate of substitution* with other things along the indifference curve that is just equal to the price paid.

Now that we have a model of consumer behavior in the form of indifference curves and budget lines, we can use this model to understand how a population's consumption behavior might adjust to change in prices, incomes or

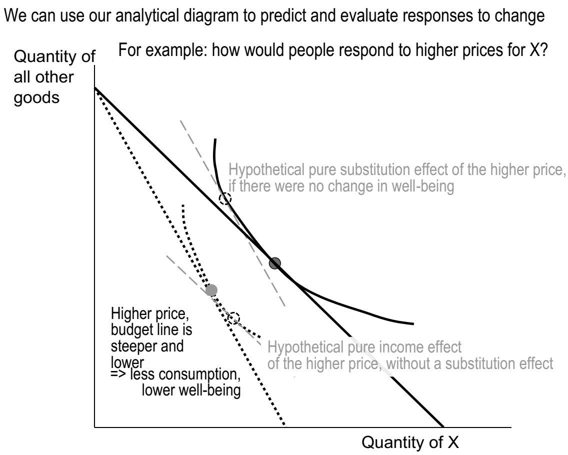


Fig. 2.7 A price increase for the X good has both substitution and income effects

preferences. In these thought experiments we will change only one thing at a time, and then combine multiple changes in more realistic scenarios. To illustrate this we show a change in prices from the solid to the dotted or dashed lines in Fig. 2.7.

Figure 2.7 shows how, when the price of food goes up, the budget line rotates inwards along the horizontal axis, towards the origin, because less of X can be purchased at the same level of income. The Y intercept stays the same, since the price of all other goods has not changed and therefore, if none of X were being consumed, the amount of other things that could be consumed is unchanged. How do we know in which direction to rotate the budget line? Remember that the slope of the budget line is the price of X, so when that increases the budget line gets steeper and the consumer can no longer reach their original level of wellbeing. They are reduced to a lower budget line, along which their best (or least bad) option is at a new point of tangency, between the new (dotted) indifference curve and the new (dotted) budget line. Remember that the lower-level indifference curve is part of the same preference mapping as the original indifference curve, so the two curves cannot cross.

The change in consumption due to a lower price is just one change but it can be understood as having two components. A first change is a reduction in the consumer's purchasing power. When prices rise, consumers cannot purchase as much as before if income stays the same. This is the *income effect* of a price change. It represents a reduction in what the person can afford. The second change is due to the new price ratio between goods. Even if the consumer were offered compensation for their loss of real income to the same level of wellbeing, they would still move along their indifferent curve because the relative price of X has changed. This is the *substitution effect* of the price change, as people adjust away from good that has become relatively more expensive.

Later in this book we will see how the framework used in these diagrams can be applied to explain, predict and evaluate the outcomes of many different changes in circumstances, including a wide range of government policies. You may want to start sketching different diagrams yourself now, to see how the logic works in various scenarios.

2.1.3 Conclusion

Nutritionists focus on measuring what people eat and how it affects their health, while economists focus on explaining and predicting changes or differences in dietary patterns. Actual events are an infinitely complicated mix of interacting forces, which economists represent as elements of each analytical diagram that distinguish between prices, incomes and preferences. In each community, the prices of available foods are likely to be similar for everyone, while incomes will differ between households and preferences will differ between individuals. In economics, we disentangle complex changes by examining one factor at a time, in a system of simultaneous equations through which everything is interconnected. So far we have seen only the drivers of food choice. In the next section we look at food production and distribution, to address actions of farmers and food sellers, before we turn to societal outcomes and government policies.

2.2 Producer Choices: Agriculture and Food Manufacturing

2.2.1 Motivation and Guiding Questions

So far, we have seen how economists explain food consumption choices. What determines food production, and how does food production interact with consumption?

In this section we analyze farming and production decisions using the same type of diagram as the previous section's analysis of food choice and consumption, building up towards a unified approach to the economics of agriculture and food systems. In this view, economists explain production choices as the best (or least bad) choice from the available options for each individual producer. We observe a bewildering variety of choices around the world, and we interpret each one as the point where a line meets a curve, at the person's highest attainable level of wellbeing. As with consumption, this framework helps explain why people do similar things when in similar circumstances, while allowing us to predict and evaluate producers' response to changes in underlying conditions and government policies.

Economists explain production with the same underlying principles as consumption, based on the observation that people have unique experiences with their own situation over time. This insight is especially important when trying to understand farmers' choices, as they are often members of multigenerational families who have farmed their lands together for decades. Farmers typically have more information about their situation, options and the consequences of each choice than any outside observer. Economists take that information into account by interpreting the actions we observe as having been chosen from among the person's limited options as the best way for them to achieve their objectives, given the difficult, often dangerous, weather-dependent and risky circumstances under which food is produced.

This textbook aims to cover all interlinked aspects of the food system, from agriculture to health. Interest in the work of farmers and food producers goes beyond their role in meeting nutrient needs. Farming is by far the most common occupation for low-income people in Africa, Asia and Latin America, and food production jobs play a similar role for many low-income people in the U.S. and other countries. These livelihoods are universally important as entry-level jobs for younger workers, as well as recent immigrants and other people who lack the formal qualifications and connections needed for employment in higher wage sectors. Farming and food production also has an outsized impact on the natural world, high vulnerability to extreme weather and climate change, and important cultural resonance as the main work for almost everyone's ancestors.

One important aspect of food systems is that farmers often consume at least some of what they produce, linking production and consumption even more directly than would be the case for other people. Another key factor is that over 90% of farms worldwide are family enterprises, owned and operated by close relatives, with almost no outside investors or salaried employees. Family farms may borrow money and rent some of the land they farm, and may hire seasonal or part-time workers, but management decisions are typically made by trusted family members. This ensures that farm sizes are typically limited by the area of land that one family can manage, whether the land is owned or rented. Only a few types of agricultural operations such as greenhouses and wineries or sugar or tea plantations attract investors and salaried managers, typically in situations where operations require less of the place-specific, weather-dependent day-to-day decision-making done by independent family farmers who live where they work.

The persistence of family farming is among the most surprising facts about the economics of food. In the U.S. and elsewhere most farms do not sell directly to consumers but operate behind the scenes, selling their produce in bulk to specialists for transport and distribution, often for use as ingredients in packaged and processed foods. Unlike farms, the food companies with whom consumers usually interact are typically owned by investors and run by hired managers. They buy ingredients from various sources, often combining produce from many different farms. Consumers everywhere in the world often seek out opportunities to buy directly from individual farmers, but that is special in part because it is relatively rare.

The reason why most farms are family owned and use mostly family labor is not because consumers prefer to buy from family farms, but because family farming is a more efficient and lower-cost way of producing most agricultural products. One underlying reason is that field crop operations require quick decisions based on location-specific information each day throughout the season. A farmer's skill and effort in planting, weed or pest control and harvesting is visible to them but very difficult for a supervisor to observe, because outcomes are heavily influenced by many intervening factors. Only someone very close to the action can distinguish skill from luck, so self-motivated family members consistently outperform hired workers.

The fact that most farms are family operations does not mean they are small in terms of land area or quantity produced. In high-income settings, farms remain in operation only if they can cover their costs and justify the management effort they require, so family operators may cultivate thousands of acres using equipment that costs several million dollars. Whether a family farm is small or large, its efficiency typically relies on workers being highly self-motivated, making efforts and making decisions based on information they observe in the fields every day. The exceptions to this rule provide important insight into the problem, as nonfamily operations tend to dominate where production is concentrated spatially and easier to supervise, such as livestock operations or sugarcane, cut flowers, and some kinds of fruit or vegetable production.

In this chapter, we will develop and use analytical diagrams to explain and predict changes in food production, to understand how production can be made more resilient, sustainable and inclusive while also meeting consumer needs for safe and nutrient dense foods in sufficient quantities for a supportive, high-quality diet. Just as our analytical diagrams for consumption began with indifference curves that are bowed in to show diminishing returns to each additional unit consumed, our diagrams in this chapter begin with production possibility frontiers that are bowed out to show diminishing returns from each additional unit produced. Those diminishing returns interacting with the relative price or value of each thing lead people to choose the quantities we observe.

You experience diminishing returns in production activities within your own life too. Think about the number of hours you might study for an exam in food economics. The first hour that you study might be hugely productive in terms of your grasp of the material. The second hour that you study would still be very productive, but not quite as productive as the first, and so on. Once you understand the concept of diminishing returns, you will start to see it everywhere.

In this chapter, you will learn how to understand farmer decisions through three difference glances into their marginal decision-making: the choice between two outputs (the *production possibilities frontier*, or *PPF*), the choice of input and output level (the *input response curve,* or *IRC*), and the choice between two inputs (the *isoquant*, or *input substitution curve*). The effects of price changes and farmer choices between these dimensions will allow us to derive supply curves and elasticity.

By the end of this section, you will be able to:

1. Describe the economic determinants of food production choices;
2. Sketch production possibilities frontiers and revenue lines, input response curves and profit lines, and isoquants and cost lines, to explain choices as points on a diagram;
3. Use the analytical diagrams to explain and predict change in agricultural production in response to change in prices, available technologies and the natural environment; and
4. Describe differences and similarities between farming and other activities in the economy.

2.2.2 Analytical Tools

The diagrams used by economists to explain production are similar to the diagrams for consumption, but in reverse. Previously we explained food choice as the point along their budget line that reaches the highest attainable indifference curve, while this section explains production as the point along a curve that reaches the highest attainable revenue or profit line. In each case, the line's slope is fixed by relative prices, explaining movements along each curve to reach a point of tangency where the curve's varying slope just equals the fixed slope of each price line. As we will quickly see, actually sketching these diagrams provides visual insights that are much clearer than any explanation in words, and are generally applicable to a wide range of specific examples.

The Production Possibilities Frontier (PPF)

In a mirror image of logic to consumer decision-making, we begin with producer choices between the quantities of two outputs: the quantity of X on the horizontal axis and the quantity of all other goods on the vertical axis. Each point on this two-dimensional diagram represents one possible choice we might observe, along a curve that shows the frontier of other production possibilities as shown in Fig. 2.8.

In Fig. 2.1 we identify the amount produced using the letter Q for a variable quantity along each axis, with a subscript to say which quantity we are talking about. In this case, along the vertical axis we show the quantity of all other goods labeled Q_a, and along the horizontal axis we show the specific product of interest that could be anything so its quantity is labeled Q_x. The combination of Q_a and Q_x we observe is the point labeled **O**, along a curve that maps out the production possibilities frontier (PPF) of all points that would be equally feasible for our farmer to grow, based on the natural conditions and technology available to her. As before, we derive this curve from the observation that farmers will do the best they can with what they have, and sketch the result in two dimensions at a given level of all other variables. With this producer's same amount of labor and other resources, the other points

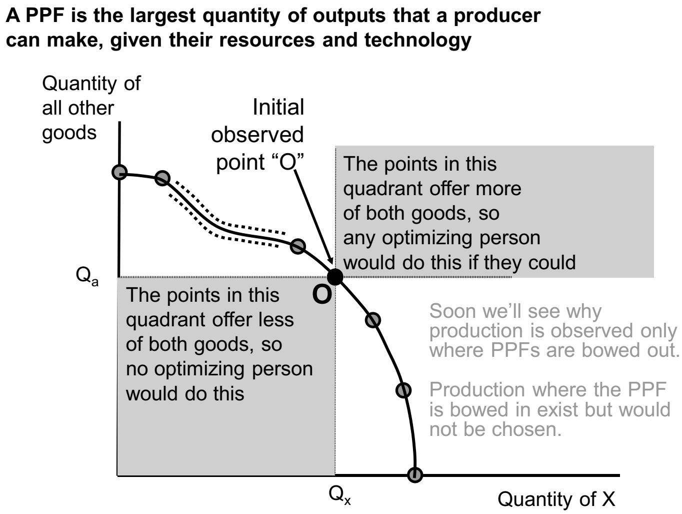


Fig. 2.8 Definition of the production possibilities frontier (PPF)

she might have chosen could not be in the top-right shaded quadrant because those would have been better and therefore chosen instead of the observed point if that were feasible for this producer, and cannot be in the lower-left shaded quadrant for the opposite reason that they produce less output and are less desirable than the observed point.

As implied by its name, the PPF is the frontier of feasible production, but unlike everyday use of the term 'frontier', economists expect all observed production to be along that curve. In other words, the frontier is defined as the feasible region for ordinary producers, who are expected to have learned from experience to do the best they can with what they have. Like an indifference curve, the PPF must be downward sloping, but in this case the curve's slope captures the incremental cost of making each additional or marginal unit of the product shown on the X axis, in terms of all other things the producer might have made with the same resources, under a given set of circumstances dictated by nature and the technologies available to this producer. As the quantity of X that she produces is increased from zero to the observed level, resources such as land and labor must have been reallocated from making other things into production of X. At some point there could be increasing returns, shown as a bowed-in portion of the PPF, where and when allocating more resources to production of X makes each additional unit more productive, but the actual observed point will be at a point along the PPF where the producer experiences diminishing returns along a bowed-out segment of the curve.

For example, in William's childhood his family kept a few chickens in a backyard shed. Going from zero to just three or four open-air scavenging chickens was very easy and took almost nothing away from other family activities such as gardening. That would yield one or two eggs each day, and feeding them grain might yield up to three eggs per day, but additional work yielded diminishing returns in terms of fewer additional eggs until the family put enough effort into properly housing, protecting and also feeding a whole flock of at least a dozen chickens. Once the shed was fenced and care practices learned, the additional work came at relatively little cost in terms of other activities and led to a yield of around ten eggs per day. Beyond that, additional efforts would again encounter diminishing returns, shown as a steeper PPF along which each incremental egg produced comes at an increasing cost in terms of other activities. As the household varied its daily egg production from zero to twenty or more, the family's PPF for eggs versus all other activities would have had some bowed-in segments, but the actual observed quantity of eggs produced was usually at a point where the PPF's curvature was bowed out or concave in shape as shown in the diagram.

Like indifference curves, the PPF is an economist's way of explaining and predicting human behavior. Producers may have explored some alternative uses of their own land and labor, but they will also have learned from neighbors and others about how best to use the resources available to them. Much of the learning process is unconscious, as people shift resources from other activities into production of X they would naturally move to the frontier of possibilities and shift along their PPF to a point of diminishing returns. In William's childhood his family kept a vegetable garden as well as the backyard chickens, and after a few years his parents had learned about the right placement and timing of operations for each type of plant. The household's PPF for vegetables, like the family PPF for eggs, had some increasing returns that made it worthwhile to take the garden seriously, with features like fencing against deer and rabbits, raised beds and a trellis for climbing beans, but also diminishing returns that limited the garden's total size to what the family could manage. Producing along the family's PPF did not require unusual skills or resources, just the typical degree of learning achieved by an average vegetable producer at that place and time. Each PPF describes the production possibilities available to a specific individual producer, but the curve's shape would be similar for other people who have the same resources and technologies available to them.

As with observed consumption along each indifference curve, explaining the producer's choice along their PPF calls for additional information about the price or value of X relative to other things. From the previous chapter we saw that a consumer's options were described by a budget line, along which she chooses the point of consumption that gives the highest attainable level of wellbeing, illustrated as the highest of many parallel indifference curves. For production, the person's options are drawn as a PPF, along which she chooses the point of production that gives the highest attainable level of income or revenue as shown in Fig. 2.9.

The straight, negatively sloped line in Fig. 2.9 is the *revenue line* of total income, showing the set of goods that a producer could obtain by exchanging X for other things in trade with other people. Starting at the observed point

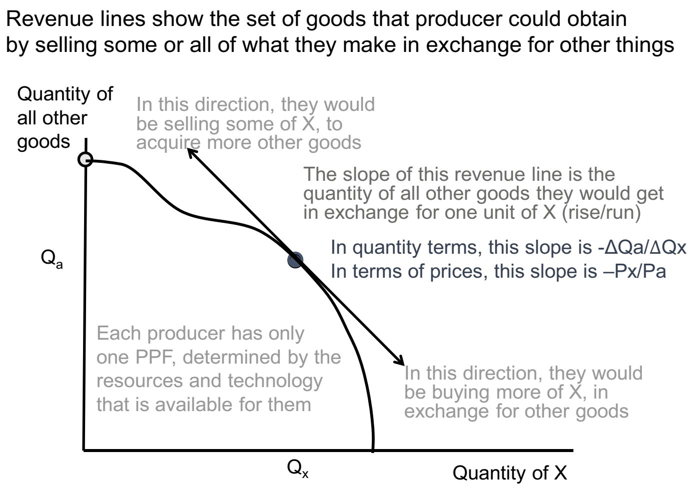


Fig. 2.9 Definition of the revenue line

of production (Q_x, Q_a), the producer might acquire more of other goods by selling some X and moving along the arrow up and to the left, or they might acquire more X than they produced by selling other things along the arrow down and to the right.

The producer's revenue line is also their income for use in consumption. The slope of that line is the rise in quantity of all other goods per unit of X that is traded with other people. If no trade with other people were possible, the producer's revenue and income would be their PPF curve itself, but when transport and storage make it possible to exchange with other people, consumption can occur along a straight line whose slope is the quantity of all other things traded for one unit of X. That slope, defined as rise/run or $-\Delta Q_a/\Delta Q_x$, is the price of X relative to all other things or $-P_x/P_a$. As with consumption, observed production is at a point of tangency along the curve where its slope just equals the price of X. The PPF's slope is the producer's *marginal rate of transformation* of all other things into production of X, while the revenue line's slope is the price of X available in trade with other people.

The PPF diagram for production, like other analytical diagrams, illustrates the fundamental principle that people have learned from experience, so when we observe their choices they have done the best they can, given what they have. This principle leads to the result that observed production is at a point of diminishing returns, where the producer's marginal rate of transformation from other things is just equal to the relative price they receive, as shown in Fig. 2.10.

Figure 2.10 shows the producer's PPF again as the curved black line, along which various possible levels of X might be produced. Straight lines whose slope is the relative price of X show levels of revenue that the producer could

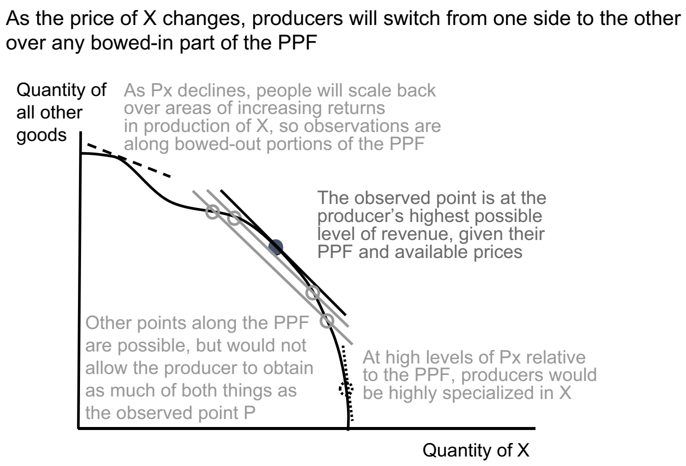


Fig. 2.10 Production we observe is each producer's choice from the options they have

obtain from each point of production. The hollow dots show various possibilities that might be observed, with the observed point being at the highest level of revenue or income that the producer's PPF would allow. As illustrated by the dashed line, at a lower price of X the producer might cut back to a lower quantity produced, potentially bypassing the bowed-in section of the PPF. Similarly at higher prices illustrated by the dotted line, the producer might increase production of X despite diminishing returns.

Over time, innovations may offer new technologies with increasing returns to additional production. The simplest kind of increasing returns comes from use of an indivisible thing like an entire machine or production method. If the relative price of X makes using or doing that thing worthwhile, producers can be expected to switch resources out of other things and use the new method up to the point where its marginal rate of transformation of other things into X is again just equal to the relative value of X compared to other things. That process is illustrated on the right side of Fig. 2.11.

Figure 2.11 shows a situation with two kinds of change in the PPF, both illustrated with no change in prices. To the right of the previously observed point, an innovation might allow farmers to adopt new equipment or other technology that offers increasing returns to greater specialization in producing more X and less of other things. To the left of the previously observed point, we show the effects of environmental degradation or climate change that reduces the production potential of this producer's resources. Both kinds of shift in producers' PPF curves occur from year to year, with growth or declines in output even when there is no change in prices. When prices change, as seen later in this book, producers would move along their PPF curves. In so doing, economists explain and predict observed points as the result of

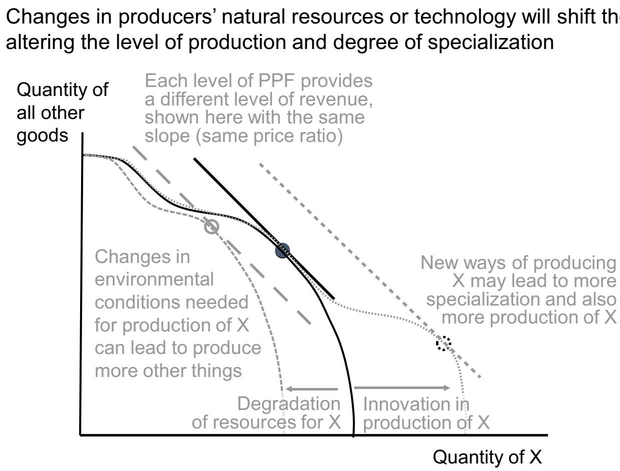


Fig. 2.11 Each producer has one PPF that shifts over time

producers having learned from experience, but any actual set of observations includes measurement error and noise or temporary adjustments to unanticipated events. It is only the average shape and location of PPFs and revenue lines that can be used for explanation and prediction with these analytical diagrams.

Each analytical diagram flattens our complex world into just two dimensions, at given levels of all other variables. The indifference curve and PPF diagrams can be drawn with the same axes as consumption decisions, with an output of interest along the horizontal axis. To complete the story, we can look at production decisions with the quantity of an input along the X axis. Economic analysis of how inputs are used in production looks somewhat similar to choices about how products are used for consumption, but there is an important difference: consumers use their income to achieve the highest attainable level of subjective wellbeing based on their own personal preferences, whereas producers use inputs to make outputs that can potentially be exchanged with other people. The options from which consumers choose are dictated by income and prices, and are shown by a budget line along which they move to reach their highest possible indifference curve. In contrast for producers, the options from which they choose are dictated by nature and technology, drawn as curves along which producers move to reach their highest level of earnings. The PPF curve explains a producer's choices between two outputs, while the curves introduced below show their choices about use of inputs.

The Input Response Curve (IRC)

Explaining a producer's use of inputs begins with the *input response curve* (IRC), showing the frontier of an output that can be produced at each level of an input. Farmers use inputs such as labor and equipment, land and fertilizer whose quantity can be shown along the horizontal axis, to produce an output whose quantity can be shown along the vertical axis as shown in Fig. 2.12.

Just like the PPF, an IRC is a frontier of technical efficiency, showing the highest possible level of output along the vertical axis that would be attainable at each point along the horizontal axis, at the same level of all other factors that might influence production such as weather and available resources. These frontiers are dictated by nature and technology available to the producer. If they have learned from experience, they would always be along these frontiers, because any higher point would be infeasible and any lower point would be undesirable.

A key fact about production that can be captured in both a PPF and an IRC is the possibility of increasing returns, highlighted in Fig. 2.12 using a dotted border around the curve. In the range of increasing returns, the IRC's slope is rising as additional inputs are applied. For example, going from zero to ten hours of labor on a strawberry field might yield zero fruit, because that is just enough time for planting and not enough time for harvesting. Reaching twenty hours might allow both planting and harvesting of some fruit, but adding another ten or more hours for weeding and pest control would make the planting and harvesting even more productive. To the extent that farmers have learned from experience they will prioritize the most important steps first, encountering diminishing returns as they add hours beyond the steepest region of the IRC.

As with the PPF, an IRC shows the producer's constraints set by nature and technology, offering a limited set of options from which to choose. Again, we expect producers to move along that curve to their preferred point, based

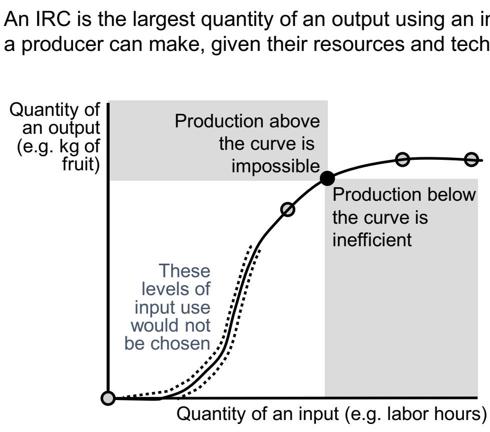


Fig. 2.12 Definition of the input response curve (IRC)

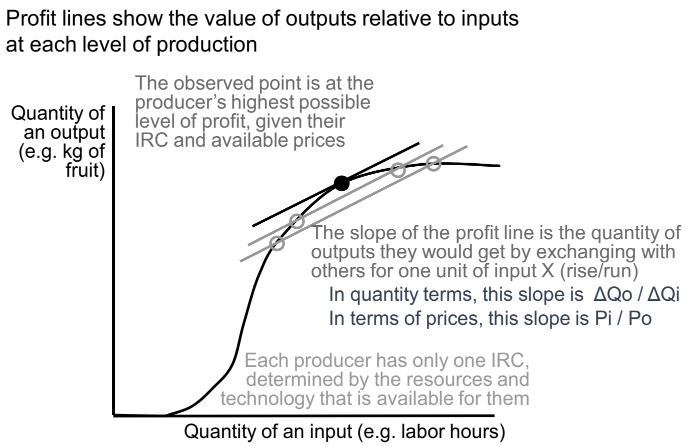


Fig. 2.13 Definition of the profit line

on the highest level of earnings they can attain. In the case of an IRC, the producer's earnings come from *profits,* defined as the value of output minus the cost of inputs, shown graphically using *profit lines* to find the most valuable level of input use as shown in Fig. 2.13.

The profit lines used with the IRC in Fig. 2.13 are similar to the revenue lines used to explain choices along the PPF in previous figures. Both are *price lines* whose slope shows the relative cost of what's on the horizontal axis, in terms of what's on the vertical axis. For example in this case, if the output were fruit that is worth $50 per bushel and the input is labor worth $10 per hour, then one bushel of fruit is worth five hours of labor, and the profit line's slope is 0.20 bu/hr ($10 per hour divided by $50 per bushel). Farmers who have learned from experience would move along their IRC until they reach the highest level of profit, at which point the IRC's slope just equals that same cost of labor in terms of fruit. Other points along the curve would all be technically efficient but are less desirable for the producer, simply because they produce a lower value of output after accounting for the value of inputs used.

The slope of each price line could reflect market prices paid or received when buying or selling, but might also reflect other costs incurred or values received. For labor use along the horizontal axis, only some farm work is paid by the hour. Most agricultural labor is done by self-employed members of a family enterprise, working to maintain their farm and earn a share of whatever the farm can produce. The family may grow barely enough to survive and avoid losing their land or other assets, but each worker would still be choosing the best of their limited options along an IRC.

Even when things are bought and sold at a market price, the economic definition of something's value is its full *opportunity cost*, referring to the best available alternative. For example each hour of family labor would be valued at that person's opportunity cost of time, including whatever else they would be doing such as caring for others or oneself. Opportunity costs vary throughout the day and among people, and may actually switch between positive and negative values. For example an activity like gardening is done by some people for enjoyment, even as others do similar work for their livelihood.

The entire opportunity cost of something that is bought or sold includes not only its market prices, but also any other *transaction costs* that must be incurred when trading with other people. Transaction costs play a large role in food systems. For example, in farm production the cost of hiring a worker is not just the wages paid but also time and effort required for supervision. Work on crop fields can be especially difficult to monitor when operations occur out of sight and affect output in ways that are not easily measured. More generally, whenever transportation or other barriers make it difficult to exchange something with others, people have to do things for themselves. When transactions are easier, people can trade with each other to provide options beyond what each person can do with their own limited resources.

The slope of each price line is set by often unknown levels of market prices, opportunity costs and transaction costs, while the shape and position of each PPF and IRC is set by highly variable environmental conditions and available technologies. Our analytical diagrams are typically impossible to quantify, but they are still very useful to provide qualitative explanations, predictions and assessments of whether, how and why outcomes might change. The lines and curves on our diagrams lead to useful insights into how people respond to change, as illustrated in Fig. 2.14.

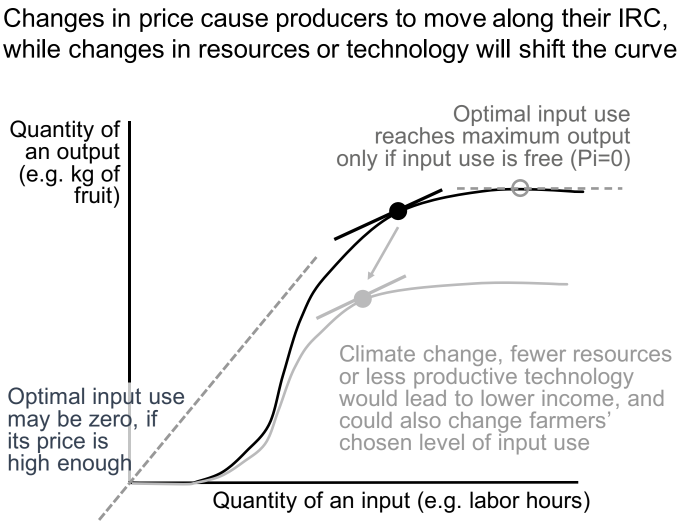


Fig. 2.14 Input use and output level will vary with prices, resources and technology

In Fig. 2.14, the initial observed point resulting from a producer's previous experience is shown as usual by the dark solid point. One kind of change would be caused by variation in prices of the output or the input, leading to movements along the same IRC. If the output becomes more valuable relative to the input, a producer would seek out higher production levels, moving up along the IRC with additional input use. The farthest extreme we could observe, if the price of the input fell to zero, is the round O at the highest possible level of output beyond which additional inputs would not add to profits. Conversely if the relative price of the input were to rise, a producer would cut back on input use, moving to the left along the IRC, and as the price line gets steeper eventually the farmer's best option would be to shut down or choose zero input use, as shown at left of the IRC. Intermediate levels of input use along the bowed-up region of the IRC would not typically be observed, because producing nothing at all would be better than that. Any production that is worthwhile would have exhausted any available increasing returns and be observed along a region of diminishing returns. This aspect of the IRC in Fig. 2.14 is similar to choices along the PPF in Fig. 2.10, which showed how producers move along their production possibilities to specialize in activities that offer economies of size or scale, up to a region where incremental changes have diminishing returns.

Another kind of change away from the observed point could be caused by nature and technology, shifting the IRC itself in ways that alter production at each price. That kind of change was presented for the PPF in Fig. 2.11, where the diagram illustrated both degradation of natural resources which reduces output at each level of input use, and also innovation towards new technologies which increase output at each level of input use. The existing PPFs and IRCs along which farmers produce is the net result of both kinds of change in the past, with some environmental harms that have reduced output and some innovations that have increased it, each of which alters the level of profits and alters farmers' decisions about input use.

The changes shown in Fig. 2.14 illustrate the consequences of climate change or other worsening of input response. These typically alter not only the level of output at each input level, but also the slope of the IRC. A worsening of input response implies a flatter as well as lower IRC at the original level of input use, shifting from the black to the gray curves. The producer would soon discover that, under their new circumstances, the old level of input use is no longer the best they can do.

As drawn in Fig. 2.14, the highest profit along the gray curves calls for a lower level of input use than before. In some cases, environmental change would make the IRC steeper at the old input level, driving producers to increase input use. Changing input levels can also be caused by innovations and new technologies that alter the IRC, including new mechanical equipment that changes the use of labor, and new agronomic techniques or biochemical inputs that can reduce fossil fuel use and hence greenhouse gas emissions.

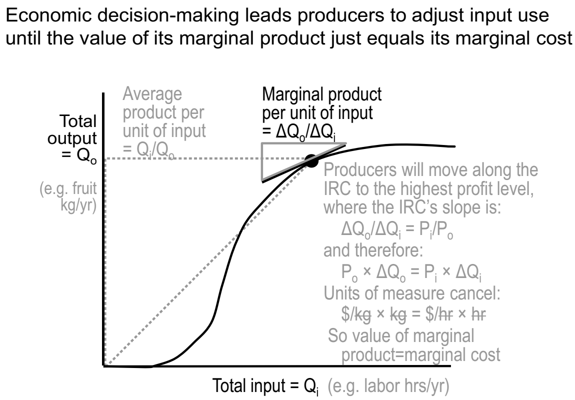


Fig. 2.15 Average versus marginal product per unit of inputs

The economics approach to explaining and predicting decisions is that choices are made based on the incremental value of each unit. The average or total value is important to see the person's level of revenue, cost or profit, but change is driven by differences in the *marginal product* of each additional unit as shown in Fig. 2.15.

The marginal product of an additional input is the IRC's slope, and at the producer's best available option that slope is also the cost of inputs in terms of the output. Figure 2.15 shows how the marginal value of an input differs from its average value. Quantities chosen are based on marginal values, yielding the average value that drives the producer's income or level of profits. Without randomized experiments an outside observer cannot observe the slope of the IRC, but economists can infer from observed behavior of producers that their expected marginal product is the marginal cost they pay for inputs. In other words, the marginal physical product of inputs along the IRC ($\frac{\Delta Q_o}{\Delta Q_i}$) would just equal the relative price paid ($\frac{P_i}{P_o}$) and is similar among producers who face similar prices, while each producer may have very different levels of average product ($\frac{Q_o}{Q_i}$) based on their resources and technology.

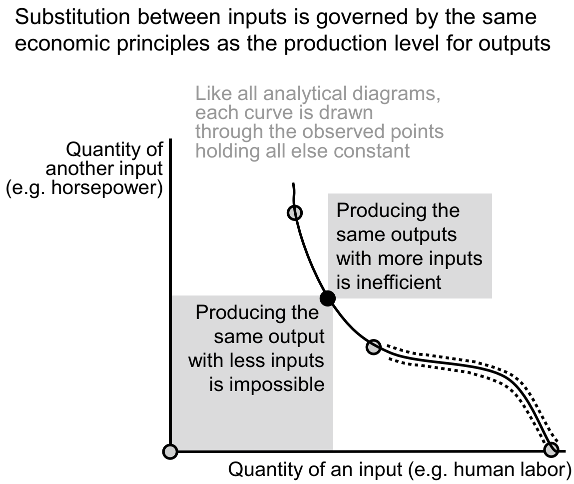
The Isoquant *or Input Substitution Curve (ISC)*

So far, we have examined producers' choice among their options for which outputs to make along a PPF, and then how much of each input to use along an IRC. The third possible way of looking at a producer's options is their choice among inputs, along an *input substitution curve* (ISC). This third view completes our set of two-dimensional diagrams showing the producer's multi-dimensional *production function*, tracing the boundaries of technical efficiency allowed by nature and available technology. The boundary on production of all outputs using all inputs can be imagined as a continuous surface, playing out all possible variations of the three curves. We could redraw these curves for every aspect of production, considering every possible pair of inputs and outputs, and all such curves would have one of the three possible shapes: either a PPF between outputs, an IRC for an input and an output, or an ISC between two inputs. In each case, economic analysis reveals how a producer might move along the curve in response to a change in prices or other circumstances towards their best available option.

Historically the curve between two inputs was called an *isoquant*, to emphasize that it traces all possible combinations for the *iso* (same) *quant*ity of all outputs. That traditional name remains in widespread use, but might be confusing in the context of this book because all two-dimensional curves in this section are all *isolines*. The PPF and IRC, like the ISC and the consumer's indifference curve, are all drawn at a constant level of all variables other than those on the two axes. Referring to the curve between two inputs as an ISC is helpful because it more specifically describes what is shown, and also complements the term IRC which shows responsiveness of output to an input. While the IRC slopes upward, the ISC or isoquant slopes downward as shown in Fig. 2.16.

To draw the ISC shown in Fig. 2.16, we can start with the observed combination of two inputs at the solid black dot. As before, if the producer has learned from experience and done the best they can at the given level of all other variables, then we can infer that it would be impossible for them to have produced the same output with less of both inputs, and undesirable or inefficient for them to have produced the same output with more of both inputs. That is why the ISC must slope down. The ISC shows the different techniques that a producer might adopt, substituting between the resource shown along the horizontal axis (such as their own labor effort, in hours of person-power per year) and the resource shown along the vertical axis (such as machinery time, in hours of horsepower or kilowatt-hours of electricity use).

Fig. 2.16 Definition of the isoquant or input substitution curve (ISC)



The downward sloping ISC could have segments that are bowed out or in. For example, as drawn in Fig. 2.16 there might be a bowed-out segment on the right when the producer first adopts some equipment instead of working entirely by hand. In this example, over the region highlighted by dotted lines, each increment of machinery up from zero offers increasing returns, working together with other mechanical parts to substitute for more and more labor as shown by a flatter slope when moving along the ISC from right to left. As with the IRC and PPF, however, that region would not actually be observed. To identify the combinations of labor and machinery that a producer might choose, we need relative prices and the resulting cost lines shown in Fig. 2.17.

Choices along an ISC are explained using the same economic principles as along the PPF and IRC, except that a producer's preferred option would have the lowest total cost, instead of the highest revenue or profit. In mathematical terms, the *cost minimization* problem shown in Fig. 2.17 mirrors the *profit maximization* used to explain levels of output, as well as *utility maximization* when consumers choose what combination of products to use in pursuit of overall wellbeing. Each diagram shows a form of *constrained optimization*, revealing the implications of people having chosen the best of their available options, as illustrated graphically in our analytical diagrams.

At this point in the text it is helpful to revisit how terms like 'optimization' have a specific meaning in economics that differs from their use in everyday life. When economists explain observed behavior as having been an *optimal choice*, shown in our diagrams as the point with the lowest available cost or the highest available revenue and profit, we are using the term 'optimal' to mean only that the action was best for that person at that time, given their options and constraints such as opportunity costs and transaction costs. In everyday use, the word 'optimal' is often used for an imagined world with fewer constraints and lower costs than in real life. Similarly, within economics

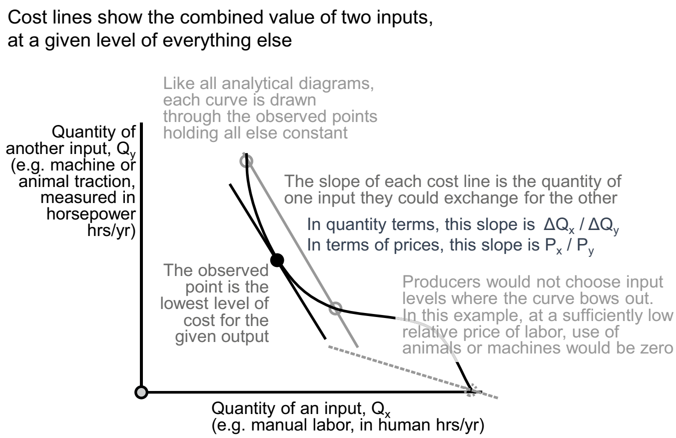


Fig. 2.17 Definition of cost lines and choice among inputs

we explain the level of things using the *marginal value* of each additional unit, and we place that quantity in the middle of our diagrams so as to explain why that was chosen instead of alternatives. In other settings, the word 'marginal' means peripheral to the main story, whereas in economics the marginal thing is central to our explanations and predictions.

In the context of Fig. 2.17, we explain the combination of inputs used by a producer as having their lowest total cost of production, shown by a *cost line* whose slope is the price of the input along the horizontal axis, divided by the price of the input along the vertical axis. That rise over run is the quantity of the input on the vertical axis that could be exchanged with other people for one more unit of the input on the horizontal axis. The available options are dictated by nature and technology, which leads to the ISC between these two inputs at a given level of all other variables. When producers have learned from experience, the best of their options is the point along that ISC with the lowest total cost.

Using the example shown in Fig. 2.17, if the price or opportunity cost incurred by the producer for each hour of labor is extremely low, production might occur with only human labor and zero animals or machinery on the right of the ISC. The cost line's slope is the relative price of labor, so if opportunities to use animals or machinery become available at lower cost per hour of work, a producer could adopt technologies that use increasing amounts of horsepower or kilowatts to replace each hour of human labor. The process of mechanization is shown here as movement along the ISC, illustrating how there is typically a region of increasing returns where adopting each additional unit of horsepower or kilowatts saves an increasing number of human labor hours. At relative prices shown by the slope of the solid cost line, mechanization offers cost reduction along the ISC only up to the observed point, due to diminishing returns that make further mechanization less attractive to the producer than their observed choice.

The economic principles that help explain technology adoption along an ISC provide important insights into how incentives guide innovation over time. When there are trends in the relative cost of things, for example rising wages compared to the cost of machinery, production can be expected to use less of the inputs that are increasingly expensive, and more of the inputs that are increasingly abundant. These trends drive the adoption of new techniques and also guide the invention of entirely new technologies, as illustrated in Fig. 2.18.

The example in Fig. 2.18 shows how a higher cost of labor, for example due to higher opportunity costs of a farmer's time or transaction costs when hiring workers, would lead producers to choose a higher level of mechanization along the solid ISC to the open circle. Furthermore, the invention of entirely new technologies could offer options to produce at even lower costs as shown by the dashed price line, saving even more labor using newly invented production methods along the dashed ISC.

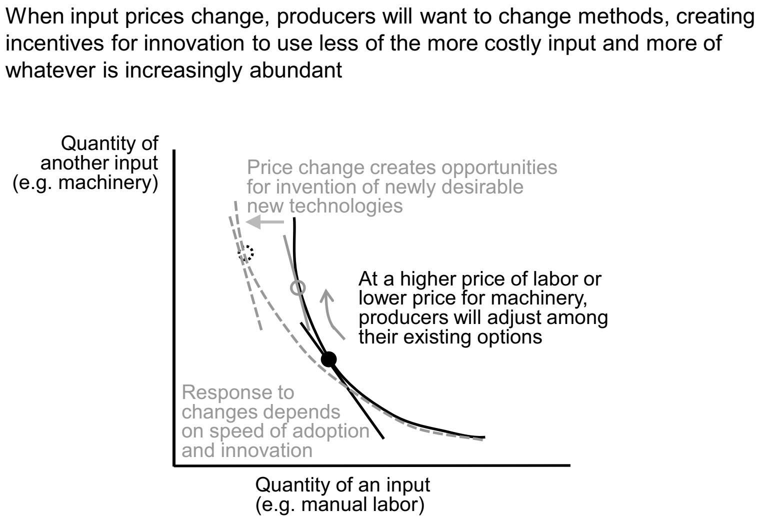


Fig. 2.18 A change in price can induce invention as well as adoption of new techniques

Differences or changes in the relative cost of inputs are sometimes predictable, driving the direction of technological change in a process known as *induced innovation*. The example of induced innovation shown in Fig. 2.18 is how higher labor costs relative to other inputs lead to not only adoption of known techniques to use less labor, but also innovations that create new options to go even further in that direction. The dashed gray price line has the same slope as the solid gray price line, but the new dashed ISC offers a flatter slope than the solid ISC at previously observed levels of labor use, leading producers to replace even more labor with machinery.

The process of induced innovation shown in Fig. 2.18 is among the most important forces affecting agriculture and food systems, driving change over time and differences between regions. Induced innovation shapes not only mechanization and employment but also use of energy and other resources. Through most of the twentieth century, steady declines in cost of fossil fuels, inorganic fertilizers and crop chemicals drove a seemingly endless trend towards use of petrochemicals, for both intensification to higher yields on existing fields and also cropland expansion. In the late twentieth century the direction of change shifted away from fossil fuels, with a rapid but not yet sufficient race towards electricity powered by renewable energy sources, and many other shifts in agriculture and food systems described in Chapters 10, 11 and 12.

2.3 Economics of Size and Scale

The means of production available at each place and time have been shown by PPF, IRC and ISC curves, tracing all possible two-dimensional perspectives on the multidimensional functions by which people could potentially convert inputs into outputs. Each production process might offer a region of increasing returns along which increasing quantities is increasingly attractive, and we expect producers to learn about those opportunities and choose options that yield the lowest available cost and highest available revenue or profit at the relative prices they face. These principles help explain why observed outcomes have diminishing returns to further changes, and also minimum and maximum quantities that are likely to be observed.

In economics, changes in the *scale* of an activity or enterprise refer to proportional changes in all inputs and other resources used. An enterprise that is 10% larger in scale would use 10% more of each thing, including 10% more hours for each type of labor as well as 10% more land and 10% more equipment and also 10% more energy. In contrast, the *size* of an enterprise refers to altering resources per worker, for example with more machinery or a larger land area. The observed scale and size of each enterprise is limited by diminishing returns to adding more of each variable input, given the enterprise's fixed factors that do not change.

The phrase *economies of scale* refers to the possibility that increasing returns to scale allow expansion to lower cost or increase revenue and profit per unit of production, while *diseconomies of scale* arise when diminishing returns impose a limit on further expansion. The intermediate case is *constant* returns to scale, where for example a proportional increase in all inputs yields that same proportional increase in all outputs. With constant returns, cost per unit is the same for enterprises of different scales.

A limiting factor determining the scale of each individual enterprise is often its management and the transaction cost of expanding operations across more different settings. For example, a given city will have various kinds of restaurants and cafeterias, each with a different number of seats and meals served per day. Owners and managers of independent restaurants serving individual customers may start with just one location, and then try to replicate or diversify their operations at different locations, but even the most successful restaurant owners and managers cannot effectively supervise more than a small fraction of all restaurants in a typical city. In contrast, large-scale institutional food service at schools, hospitals and other facilities is more suited to centralized management, so cities may have just a few big commercial food service providers.

The economics of size and scale concerns both the magnitude of each individual enterprise and also the cost per unit sold for an entire sector of production. Management challenges limit the size and scale of individual enterprises, but an ecosystem of many enterprises can often expand with constant or even increasing returns to scale until the whole sector encounters its own diminishing returns. For example, a city might have a wide range of restaurants that serve various meals at different prices, and that entire ecosystem of restaurants might expand or shrink over time. Enterprises often benefit from each other's presence, leading to agglomeration in geographic clusters of similar activities. The benefits of agglomeration are often visible in the restaurant sector, as establishments choose to locate near each other and neighborhoods with many similar restaurants often have higher quality and lower prices. Agglomeration effects occur between sectors as well. The initial start of a cluster may be influenced by transportation routes or other geographic factors, but then various kinds of activities will benefit from proximity to each other, leading to urbanization and the growth of each individual town or city even as the surrounding rural area remains cultivated by dispersed family farmers in rural areas.

Scale economies for individual enterprises and for entire sectors play a crucial role in agriculture and food systems, determining the size and structure of organizations that can sustainably undertake each kind of activity. The smallest restaurants we typically see have enough tables or take-out business to keep several people busy for much of the day. It may be operated by an owner who lives near the premises, but almost all restaurants have multiple employees and many are run by salaried managers. In contrast, most farms have zero salaried employees, even in the U.S. or other industrialized countries. Most farms are owned and operated by family members who live on site, often hiring part-time workers only for specific operations where supervision and transaction costs are low. Year-round employees are observed mainly in concentrated livestock operations, production of fruits and vegetables, or crops that require on-farm processing such as sugar or tea, where there are scale economies derived from equipment and facilities and tasks for which workers can be hired and supervised relatively easily. As technologies change, the number of workers as well as the area of land or number of animals that can effectively be managed in each individual operation, as well as the number of such operations in each area, changes with shifts in production technology and relative prices.

The enterprises we actually observe in each part of the food system are big enough to have survived, somewhere between the minimum and maximum size of feasible operation for each activity. A helpful way to describe economies of size and scale is to distinguish between *fixed costs* of big, lumpy or indivisible capital investments, in contrast to *variable costs* of applying increasing quantities of a continuous input. Fixed costs include buildings and facilities as well as management skills and other assets that are specific to an enterprise but can be used repeatedly over time, while variable costs include all materials and other inputs that are used up in production. Fixed costs are often the source of increasing returns that determine the minimum scale typically observed, while variable costs often encounter diminishing returns that limit the size of each operation.

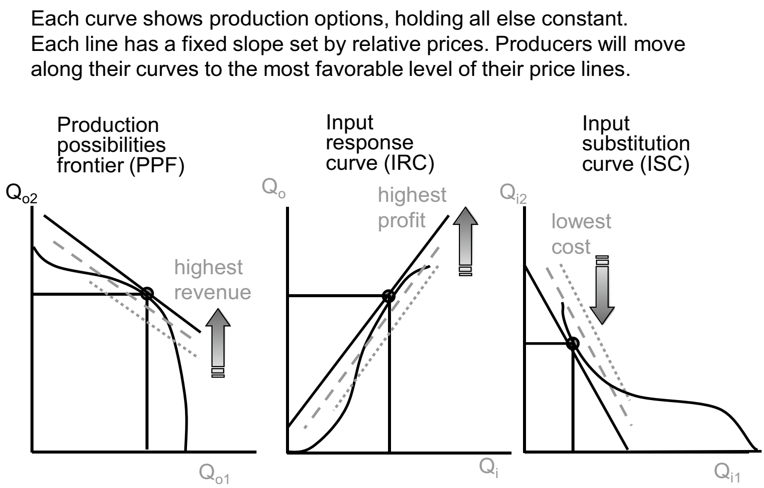


Fig. 2.19 Summary of all three two-dimensional perspectives on production

To explain the size and type of operations we are likely to see at each place and time, it is helpful to keep all three of our production diagrams in mind as shown in Fig. 2.19.

The trio of analytical diagrams in Fig. 2.19 shows how the observed point of production is chosen as the best option for that producer, offering their highest level of revenue or profit and also their lowest cost per unit of output. The slope of each price line is the relative value or cost of incremental units along the curve, where prices include all opportunity costs and transaction costs of transaction with other people. Meanwhile the shape and position of each curve are dictated by nature and technology, embodying all past investments that determine what can be made with additional inputs at each place and time.

More advanced classes in economics represent production choices mathematically using multivariate calculus and real analysis, generalizing the graphical approach illustrated in our two-dimensional diagrams. Advanced methods are helpful to explore special cases and details not covered in this introductory textbook, but the principles of economics can readily be summarized as the consequences of people having chosen the best of their available options. Redrawing these diagrams around any given decision will reveal how these principles play out in each situation, as producers choose among inputs to obtain outputs. The diagrams could be redrawn for specific people making particular things, using concrete numbers of each input and output, but the important thing is to recall the definition of each line and curve in terms of the variables shown in each axis. Once you have practiced sketching these diagrams, starting with the axes then curves and lines leading to the observed points, you will see that there is need to memorize examples because you can always redraw a new diagram for each situation.

A key feature of our individual-choice diagrams in this chapter is that the axes show quantities, measured in natural units of something such as weight, volume or servings of food, land area and labor time or energy use. Prices are used here only in relative terms, showing the relative value or cost of each thing when exchanging it for other things. The diagrams used in this chapter can help explain individual choices in food system decisions that may not involve any market transactions at all, as shown in Fig. 2.20.

The diagrams in Fig. 2.20 begin our analysis of the entire food system, showing the interaction between production and consumption for an individual person. The diagram allows us to imagine the choices of a farmer who is entirely self-sufficient, and does not exchange anything at all with other people. The diagram focuses on one of their foods they grow and eat, for example beans. Their production options between beans and all other things are limited by their PPF, along which the highest level of wellbeing is at the hollow O based on their consumption preferences shown by the dotted indifference curve. Other points along their PPF are equally possible but would be less preferred in consumption. The left side diagram shows this farmer in a situation where other people offer to buy beans from them in exchange for other things, while the right diagram shows a situation where other people offer to sell beans to them in exchange for other things.

Starting with the left diagram in Fig. 2.20, if other people offer to buy some beans along a steeper price line than the slope of the farmer's PPF at their self-sufficient level of production, the farmer could reach a higher level

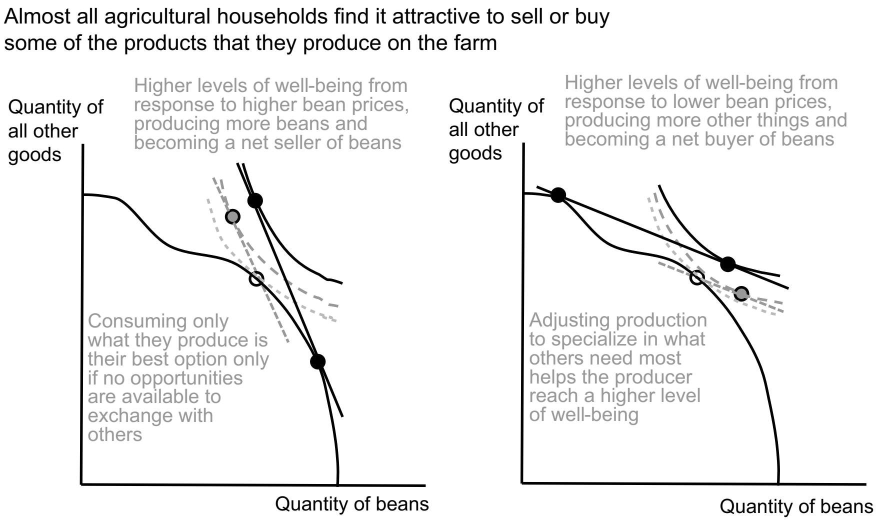


Fig. 2.20 Production and consumption for the farming household

of wellbeing by selling some of the beans they produced leftward along the price line up to the gray dot which reaches the dashed indifference curve. Learning from experience, however, the farmer would soon discover that they can reach even higher wellbeing by moving production along their PPF to the right, increasing production of beans so as to sell a larger quantity and reach the solid indifference curve which turns out to be the best of their available options, given their production options and consumption preferences.

Now turning to the right side of Fig. 2.20, we see the identical farmer in a situation where other people offer to sell them some beans at a lower price than the slope of their PPF in self-sufficiency. Again we can see that the farm could improve their wellbeing by accepting the offer, selling some of their beans from the original point of production rightward along the price line down to the gray dot which reaches the dashed indifference curve. Again, however, we would expect them to learn from experience, and soon discover that they can reach an even higher level of wellbeing by moving production along their PPF to the left, reducing production of beans so as to make more other things which they sell to others and reach the solid indifference curve, which in this case is the highest they can reach given their production options and consumption preferences.

Taken literally, the diagram refers to an individual farmer living alone, but we can also use the diagram to describe a farm household that pools their resources and makes joint decisions in service of the whole family's wellbeing. In later chapters we will address some of the ways in which households do not act like individuals, for example due to differences between household members in their preferences and bargaining power. Gender and age disparities within each household can be extremely important for nutrition and health, and for the wellbeing for women and children generally. We will return to that topic but for now we can imagine the benchmark case of a unified household that is either one individual or a family that acts together as if they were a single farmer who consumes some or all of what they grow.

Comparing the two sides of Fig. 2.20 is the foundational discovery of economics, showing how exchanging goods with other people helps each person or joint household reach a higher level of wellbeing for themselves and their children. The magnitude of gain depends on the details of each line and curve, but the qualitative discovery is that gains from trade exist whether other people want to buy from us or sell to us. In either case, remaining entirely self-sufficient is possible but undesirable and therefore unlikely to be observed. Exchanging with others, whether buying or selling, helps farmers overcome diminishing returns on their own farm in both production and consumption. This observation helps explain why even the most ancient archeological sites show evidence of food trade, and even the most remote people who value self-reliance choose to exchange with other people at least some of what they produce and consume.

The final analytical diagram to complete this chapter shows how economists can use PPFs and indifference curves for an individual farmer to explain and predict response to change. Students can redraw these diagrams for any imaginable scenario, identifying cause and effect for changes in nature or technology and hence production possibilities, changes in market conditions and hence relative prices, or changes in preferences and hence the shape of each indifference curve. The example shown is the impact of a lower relative price of beans than was used to draw the farmer's previous choice, as illustrated in Fig. 2.21.

The impact of a lower price of beans on the farmer's wellbeing depends on whether they are buying or selling beans to other people. As shown in Fig. 2.21, the farmer is always producing and consuming some beans, with the left diagram showing a *net seller* who produces more than they consume, and the right diagram showing a *net buyer* who consumes more than they produce. In this picture, the only reason for the difference is what others are willing to do. The left diagram shows a net seller because others have offered to buy their beans at a relatively high price, and the right diagram shows a net buyer because others have offered to sell them beans at a relatively low price. For the net seller, a lower price of beans reduces the gains from trade and lowers their wellbeing, as shown by the switch to the dashed price line, gray dots and dashed indifference curve. For the net buyer, a lower price of beans increases the gains from trade and raises their wellbeing.

Figure 2.21 clearly reveals how the initial direction of trade drives our qualitative conclusions about the direction of cause and effect, while the shapes of each curve influence magnitudes. On the right diagram, the initially low price of beans had led this farmer to specialize in other things, and the quantity of beans they produce is not much affected by further reduction in market price.

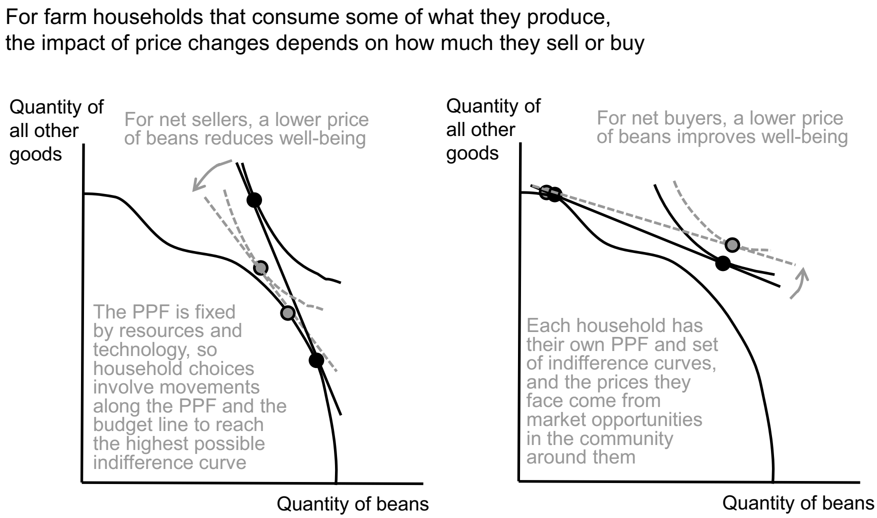


Fig. 2.21 Impact of a lower price on net sellers and net buyers

For example, they might have just a small backyard garden, and the lower price of beans allows them to buy more beans and also spend more money on other things. Meanwhile the left diagram showed the farmer putting more variable costs into moving along their PPF towards production of beans for sale to others, and the lower price leads them to cut back on that. In either case the farmer's consumption preferences is such that the quantity of beans consumed changes relatively little, and the price alters wellbeing mostly through income available to consume other things.

2.3.1 Conclusion

This long chapter spells out the economic principles used to explain and predict changes in an individual person's choices for production and consumption. Our analytical diagrams reveal how the quantities we observe being produced and consumed are the result of choices, as each person selected actions to meet their needs given their options. This approach focuses our attention on understanding and improving those options. We also recognize that some aspects of behavior may have been random and unpredictable, or preordained and unchangeable. Our focus is on the kind of behavior that was described by Alfred Marshall in 1890 at the start of his *Principles of Economics* textbook as 'the ordinary business of life', regarding 'the material requisites of wellbeing'. The underlying first principle of economics, underlying all else in this textbook and other work in economics, is that people might have chosen what we observe because it was the best of their options. The result of each person's everyday choices can be sketched graphically in two dimensions, leading to a set of causal models that make clear predictions about how people will respond to a change in production possibilities, prices and income, or preferences. The resulting theory of change is an abstract simplification of the infinitely complex world, but it sets economists on a profoundly human journey of exploration to understand and improve societal outcomes.

By design, economics is not a single complete theory of everything, but a way to create customized models suited to answering various questions about everyday living standards. Each analytical diagram is a different model, suited to different circumstances and scales of observation. Our goal in this textbook is to spell out a toolkit of interconnected models used in the economics of food, linking agriculture and natural resource use to human nutrition and health. This chapter provides a first set of modeling tools, using analytical diagrams to explain and predict individual choices in food consumption and production, as people learn from experience and move among their available options along each line or curve towards their preferred choice shown by the observed point. In the next chapter we connect the dots between each person's choices to explain and predict societal outcomes, meaning the prices and quantities produced or consumed by an entire population.