Lecture Notes[1]

ECON 323
Microeconomic Theory

Guoqiang TIAN
Department of Economics
Texas A&M University
College Station, Texas 77843
(gtian@tamu.edu)

This version: February, 2023

[1] This lecture notes are for the purpose of my teaching and convenience of my students in class. Please not distribute it.

Contents

Part I

Preliminaries and the Basics of Demand and Supply

Part 1 surveys the scopes of economics and microeconomics as well as introduces basic concepts and methodologies. We then discuss the market model of demand and supply.

Chapter 1

Economics Review and Math Review

1.1 Economics Review

Vernon Ruttan (1924–2008), a development economist at the University of Minnesota, once made the following remark:

> Economics is hard to teach. Students want to learn about the economy; instructors want to teach about economics.

This conflict is largely because there is no such thing as ”economy”, and instead, there are only theories (models) of the economy. Indeed, what we ordinarily view as a fact turns out to be a predictive statement, which is what we mean by a theory.

This course is also harder to learn since it involves more technical methods and logic thinking, which is somewhat challenging to many of our students.

1.1.1 The Themes of Economics

Economics studies economic phenomena and the economic behavior of individual agents — consumers, workers, firms, government, and other economic units as well as how they make ***trade-off choices*** so that limited resources are allocated among competing uses.

Because of the ***fundamental inconsistence and conflict*** between **limited resource** and individuals' **unlimited desires** (i.e., wants), economics is created to study trade-off choices in resource allocation to make the best use of limited resources in maximizing the satisfaction of individuals' needs.

Typical Resources

- Labor - includes the mental and physical skills provided by workers, e.g., teachers, managers, dancers, and steel workers.
- Capital - includes man-made aids to production e.g., machinery, buildings, and tools.
- Land - all natural resources, e.g., soil, forest, minerals, and water.

Two Most Objective Realities

Economic issues are difficult to solve due to the following two most objective realities:

(1) **self-interest behavior**: individuals (in whatever level: nation, firm, household, or personal level) usually pursue their own interests under normal circumstances;

(2) **asymmetric information**: information among individuals is often asymmetric, and it is easy to pretend or lie.

How to deal with these two most **basic objective realities**, what kind of economic system, institution, incentive mechanism and policy should be used have become the core issues and themes in all areas of economics.

At the same time, economics often involves **subjective value judgments**, which make economic issues even hard to be solved since it is hard to have consensus.

Strong Externality in Application

The practical application of economics has strong externality, either positive or negative. Poor applications of economic theories may affect all aspects of individuals and even a whole economy. As such, learning well and correctly understanding economics, especially the basic contents of microeconomic theory in this course are very important for practical applications.

Four Basic Questions That Must Be Answered by Any Institution

i) What goods and services should be produced and in what quantity? (Chapters 1-5)
ii) How should these goods and services be produced? (Chapters 6-14)
iii) For whom should they be produced and how should they be distributed? (welfare analysis: Chapters 9 and 11)
iv) Who will make decisions and by what process? (economic system: Chapter 1)

These questions must be answered in all economic systems, but different economic institutional arrangements provide different answers.

Whether an institutional arrangement can effectively resolve these problems depends on whether it can properly deal with ***four*** key words: **self-interests**, **information**, **incentives**, and **efficiency**.

Two Basic Economic Institutions Used in Practice:

(1) **Centrally planned (command) economic institution** (mainly a ***centralized*** decision system):

- all of the four questions are answered by the government, who determines most economic activities and monopolizes decision-making processes and all sectors;

- no unemployment, no inflation, and no free enterprises.

(2) **Market economy institution** (mainly a ***decentralized*** decision system):

- consumption and production decisions are made through market mechanisms;
- consumers and producers are motivated by pursuing self-interests.

While a real-world economic system is somewhere in between these two extremes, the key is which one is in the dominant position. The fundamental flaw of the centrally planned economic system is that it cannot effectively resolve the problems induced by **information** and **incentives**, which in turn results in **inefficiency** in allocating resources.

The market economic system has been proved to be only economic institution so far that can keep an economy with sustainable development and growth. It is the most important economic institution discovered for reaching cooperation and solving the conflicts among individuals. Modern economics studies various economic phenomena and behaviors under market economic environment by using an analytical approach, such as the demand and supply model. ***This is why this course fully focuses on the market system and will discuss how a market works***.

Methodology of Scientific Economic Analysis

Scientific economic analysis, especially aimed at studying and solving major practical problems affecting the overall situation, is inseparable from the scientific method of **three dimensions and six natures**:

(1) **Three dimensions**: theoretical logic, practical knowledge, and historical vision.

(2) **Six natures**: scientific, rigorous, realistic, pertinent, forward-looking, and thought-provoking.

Studying and especially solving major social economic problems need not only theoretical logic analysis and empirical tests, but also the confirmation of historical experiences.

Merely theory and practice may not enough, causing shortsightedness. The short-term optimum is not necessarily the long-term optimum. Historical confirmation and comparison from a broad perspective are also requisite for gaining experience and drawing lessons. Of course, if merely relying on historical experience, men's cognition will be deficient and stick in the mud, it is difficult to have innovation and creation, and will hinder economic and social development.

Only through the three dimensions of "theoretical logic, practical knowledge, and historical vision", can we guarantee that a solution or reform measure satisfies the "six natures".

1.1.2 The Themes of Microeconomics

Microeconomics is the core of economics and the theoretical foundation of all branches of economics and business science. It enables us to employ simplified assumptions for in-depth analyses of various aspects of the complex world in order to get some useful insights.

Microeconomics is the study of economic behaviors of individuals such as consumers, firms, workers and investors, as well as how markets are operated.

Macroeconomics, on the other hand, is the study of a national economy as a whole.

The core of microeconomics is **pricing**. It focuses on such questions as: how pricing is determined? which factors affect pricing? Does a firm have market (pricing) power? How can an enterprise get the power of pricing? To elucidate the answers to such a large issue, it is necessary to study the demand, supply, characteristics and functions of the market, and pricing in all kinds of markets and economic environments. As a result, microeconomics is also known as the **price theory**.

Theories and Models

Like any science, *economics is concerned with the **explanations** of observed phenomena and **predictions** based on theories*.

Theory: Developed from a ***set of hypotheses***. It is a rational type of abstract thinking about a phenomenon, or the results of such thinking.

Economic model: A simplified, often ***mathematical***, framework, which is based on economic theory and is designed to illustrate complex processes and make predictions.

When evaluating a theory, it is important to keep in mind that it is invariably imperfect and has limited success in making predictions. Why?

The two tenets of pragmatism/instrumentalism:

1) there are no complete facts, only theories (models);

2) theories are to judged by their usefulness (just like tools).

How to model an economic phenomena or issue is not only science but also art. Statistics and econometrics enable us to measure the ***accuracy*** of our predictions while historical experiences can be used to confirm the correctness of explanations and predictions from a theory.

Two Categories of Economic Theory

Economic theory can be further divided into two categories:

- **Benchmark theories**: It is for providing various benchmarks and developing reference systems, which provides necessary criteria for judging what is better and whether it constitutes the right direction.

  When one tackles a problem, it is necessary to first determine what to do, or provide the direction and goals of improvement towards the ideal situations such as **perfect competition**. Only learning from the best and comparing with the best, can we perpetually get better and even better.

The first three parts (Chapters 2-9) of this course provide such benchmark theories.

- **More realistic theories**: That aims to solve practical issues, so that assumptions are closer to reality, which are usually modifications to the benchmark theory. Parts V provides such theories.

Major Economic Agents

- consumers who generate demands for goods and services via ***utility maximization***;

- producers (firms) who supply quantities of goods and services via ***profit maximization*** (or ***loss minimization***).

Three Basic Assumptions

1. **Self-interest behavior**: individuals are self-interested and pursue their personal goals.
2. **Rationality assumption**: market participants make rational (i.e., optimal) decisions.
3. **Scarcity of resources**: market participants confront scarce resources.

1.1.3 Positive versus Normative Analysis

Positive questions deal with ***explanation and prediction****,* while *normative questions deal with what* ***ought*** *to be*:

- **Positive statements**: Describing relationships of cause and effect. Tell us what is, was, or will be. Any disputes can be settled by looking at facts, e.g., "the sun will rise in the east tomorrow".
- **Normative statements**: Opinions or value judgments; those tell us what should or ought to be. Disputes can **not** be settled by looking at facts, e.g., "It would be better to have low unemployment than to have low inflation."

Positive analysis is central to microeconomics. Normative analysis is often supplemented by value judgments. When value judgments are involved, microeconomics cannot tell us what the best policy is. However, it can clarify the trade-offs and thereby help to illuminate the issues and sharpen the debate.

1.1.4 What is a Market?

Market: The market constitutes a modality of trade in which buyers and sellers conduct voluntary exchanges. It refers not only to the location where buyers and sellers conduct exchanges, but also to all forms of trading activity, such as auction and bargaining mechanisms.

It is crucial to keep in mind that any transaction in *the market has both buyers and sellers*. In other words, for a buyer of any good, there is a corresponding seller. The final outcome of the market process is determined by the rivalry of relative forces of sellers and buyers in the market.

An **industry** includes only sellers but not buyers.

Competitive versus Noncompetitive Markets

Perfectly competitive market: market with many buyers and sellers, so that no single buyer or seller has an impact on price. That is, ***everyone takes the price as given***.

Although perfect competition is an ideal situation, many other markets are competitive enough to be treated as if they were perfectly competitive.

Other markets containing a small number of producers may still be treated as competitive for purposes of analysis. Some markets contain many producers but are **noncompetitive**; that is, ***individual firms can jointly affect the price***.

In markets that are not perfectly competitive, different firms might charge different prices for the same product. The market prices of most goods will fluctuate over time, and for many goods the fluctuations can be rapid. This is particularly true for goods sold in competitive markets.

1.1.5 Real versus Nominal Prices

Inflation is an increase in the level of prices of the goods and services that households buy, which is measured as the *rate of change of those prices*.

These are two kinds of prices:

- **Nominal price**, also known as the **absolute price**: the price unadjusted for inflation.

  For example, the nominal prices of a pound of butter was about \$0.87 in 1970, \$1.88 in 1980, \$1.99 in 1990, \$3.48 in 2015.

- **Real price**, also known as the **relative price**: the price adjusted for inflation.

  For example, the price of the first-class ticket for Titanic in 1912 was \$7,500 which is the equivalent of roughly \$128,000 in 2016 dollars.

Moreover, we have:

- **Consumer Price Index (CPI)**: measure of the aggregate price level.
  - Records the prices of a large market basket of goods purchased by a "typical" consumer over time. When the CPI is higher this year than last, it means there has been inflation since last year.

  - ***Percent changes in CPI measure the rate of inflation.***

- **Producer Price Index (PPI)**: measure of the aggregate price level for intermediate products and wholesale goods.

The Formula for Computing the Real Prices of a Good

The real prices of a good in year t in term of the base year dollars is calculated as follows:

$$\text{Real price in year } t = \frac{CPI_{\text{base year}}}{CPI_{\text{year } t}} \times \text{the nominal price in year } t.$$

Example 1.1 (Real price of butter) The nominal prices of a pound of butter was about \$0.87 in 1970, \$1.88 in 1980, \$1.99 in 1990, \$3.48 in 2015. The CPI was 38.8 in 1970 and rose to about 237 in 2015.

After correcting for inflation, was the price of butter more expensive in 2015 than in 1970?

To find out, let's calculate the 2015 price \$3.48 of butter in terms of 1970 dollars.

$$\text{Real price in 2015} = \frac{CPI_{1970}}{CPI_{2015}} \times \text{the nominal price in 2015} = \frac{38.8}{237}\$3.48 = \$.57.$$

The nominal price of butter went up by about 300 percent, while the CPI went up 511 percent. Relative to the aggregate price level, butter prices fell.

Example 1.2 (The price of eggs and the price of a college education) Table 1.1 in shows the nominal price of eggs, the normal cost of a college education, and the CPI.

Using the formula for computing the real price of a good, the real price of eggs for the year 1980 and a college education for the year 1990 in 1970

	1970	1980	1990	2000	2016
CPI	38.8	82.4	130.7	172.2	241.7
Nominal Prices					
Eggs	\$.61	\$.84	\$1.01	\$.91	\$2.47
Education	\$1,784	\$3,499	\$7,602	\$12,922	\$25,694
Real Prices (\$1970)					
Eggs	\$.61	\$.40	\$.30	\$.21	\$.40
Education	\$1,784	\$1,624	\$2,239	\$2,912	\$4,125

Table 1.1: The real prices of eggs and a college education in 1970 dollars

dollars can be found as follows:

$$\begin{aligned}\text{Real price of eggs in 1980} &= \frac{CPI_{1970}}{CPI_{1980}} \times \text{the nominal price in 1980} \\ &= \frac{38.8}{82.4}\$.84 = \$.4,\end{aligned}$$

$$\begin{aligned}\text{Real price of eggs in 1990} &= \frac{CPI_{1970}}{CPI_{1990}} \times \text{the nominal price in 1990} \\ &= \frac{38.8}{130.7}\$1.01 = \$.3,\end{aligned}$$

$$\begin{aligned}\text{Real price of education in 1980} &= \frac{CPI_{1970}}{CPI_{1980}} \times \text{the nominal price in 1980} \\ &= \frac{38.8}{82.4}\$3,499 = \$1,624,\end{aligned}$$

$$\begin{aligned}\text{Real price of education in 1990} &= \frac{CPI_{1970}}{CPI_{1990}} \times \text{the nominal price in 1990} \\ &= \frac{38.8}{130.7}\$77,603 = \$2,239,\end{aligned}$$

and so forth.

If we want to convert the CPI in 2000 into 100 and determine the real prices of eggs and a college education in 2000 dollars, we need to divide the CPI for each year by the CPI for 2000 and multiply that result by 100, and then use the new CPI numbers in Table 2.3 to find the real price of butter in 2000 dollars.

	1970	1980	1990	2000	2016
CPI	22.5	47.9	75.9	100	140.3

Table 1.2: Conversion of the CPI in 2000 dollars

Other examples on real prices are:

- In nominal terms, the minimum wage (currently \$7.25 per hour, Texas adopts the federal minimum wage rate) has increased steadily over the past 80 years. However, in real terms its 2020 level is around 33 percent lower than the minimum wage in 1970.

- The prices of both health care and college textbooks have been rising much faster than overall inflation. This is especially true of college textbook prices, which have increased about three times as fast as the CPI.

1.1.6 Why study microeconomics?

Microeconomic concepts are used by everyone to assist them in making choices as consumers and producers. The following examples show the numerous levels of microeconomic questions necessary in many decisions.

Example 1.3 (Corporate Decision Making) For instance, the Toyota Prius. In 1997, Toyota Motor Corporation introduced the Prius in Japan, and s-tarted selling it worldwide in 2001. The design and production of the Prius involved not only some impressive engineering (using hybrid power system), but a lot of economics as well. Many of the challenging questions can be answered by microeconomic theory:

- How strong in demand and how quickly will it grow? One must understand consumer preferences and trade-offs (Chapters 3-5).
- What are the production and costs of manufacturing? (Chapters 6-7)
- Given all costs of production, how many should be produced each year? (Chapters 8-9)
- Toyota had to develop pricing strategy and determine competitors reactions? (Chapters 10-11)
- Risk analysis. Uncertainty of future prices: gas, wages. (Chapter 5)
- Organizational decisions — integration of all divisions of production. (Chapters 12-13)
- Government regulation: Emissions standards (Chapters 2 and 9)

Example 1.4 (Public Policy Design) For instance, 1970 Clean Air Act imposed emissions standards and have become increasingly stringent. A number of important decisions have to be made when imposing emissions standards:

- What are the impacts on consumers?
- What are the impacts on producers?
- How should the standards be enforced?
- What are the benefits and costs?

1.2 Math Review

In this course, we focus on functions where independent and dependent variables are **real numbers**. The set of all real numbers is denoted by $\mathcal{R}$. A function of one variable, f, is a **mapping** from the **domain** X (a subset of $\mathcal{R}$) to the **range** $\mathcal{R}$, such that to every element in the domain, f assigns a unique element from the range. In short we write: $f : X \to \mathcal{R}$. Sometimes, we also write $y = f(x)$. In this case, we say x is the **independent variable** and y is the **dependent variable**.

A function of one variable can be visualized:

- The horizontal axis (often called x-axis) represents the domain.
- The vertical axis (or y-axis) represents the range.
- A generic point on the curve has a coordinate $(x, f(x))$.

1.2.1 Equations for straight lines

Equation for a straight line (or a linear function of one variable) takes the following form:

$$y = ax + b$$

where the constant a is the slope, which describes the rate at which function value changes with respect to the change of the independent variable x, and the constants b and $-\frac{b}{a}$ are respectively the y-axis and x-axis intercepts, which describe where the line intersects with the y-axis and x-axis, respectively.

Moreover, linear functions could exhibit the following geometrical features:

- positive slope ($a > 0$) - upward sloping;
- negative slope ($a < 0$) - downward sloping;
- larger absolute value in slope - steeper;

- smaller absolute value in slope - flatter;

- positive y-intercept - intersect with y-axis above x-axis;

- negative y-intercept - intersect with y-axis below x-axis;

- increase b - upward shift of the line;

- decrease b - downward shift of the line.

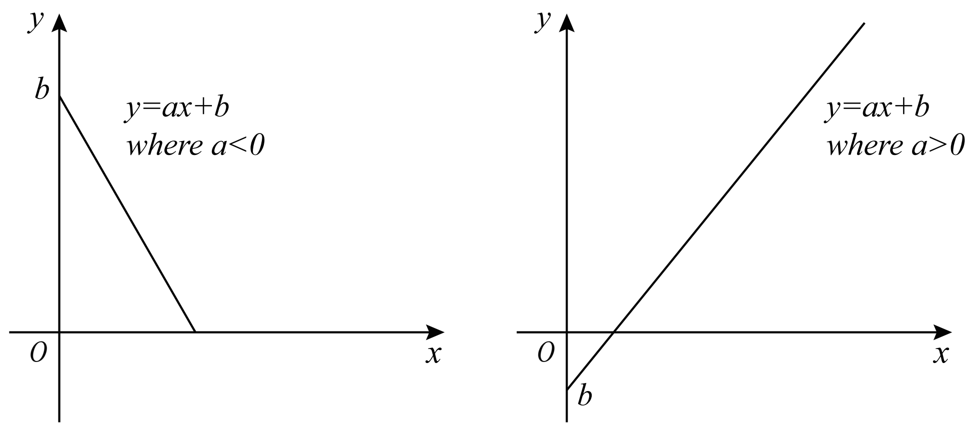

Figure 1.1 Linear functions with negative or positive slopes

1.2.2 Solve two equations with two unknowns

Example 1.5 $y = 60 - 3x$.

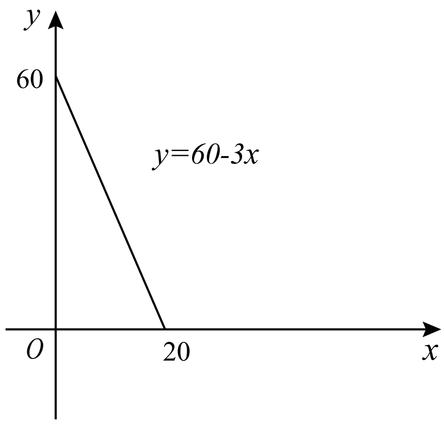

Figure 1.2 A numerical example of downward-sloping linear functions

The slope $= -3$.
$x = 0 \Rightarrow y$-axis intercept $= 60$.
$y = 0 \Rightarrow x$-axis intercept $x = 20$.

Example 1.6 $y = 5 + 2x$
The slope $= 2$.
$x = 0 \Rightarrow y$-axis intercept 5.
$y = 0 \Rightarrow x$-axis intercept $x = -\frac{5}{2}$.

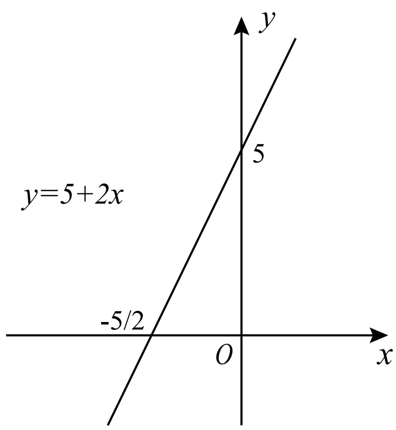

Figure 1.3 A numerical example of upward-sloping linear functions

Example 1.7 To find the intersection of these two lines, we have the following two approaches:

Graphically:

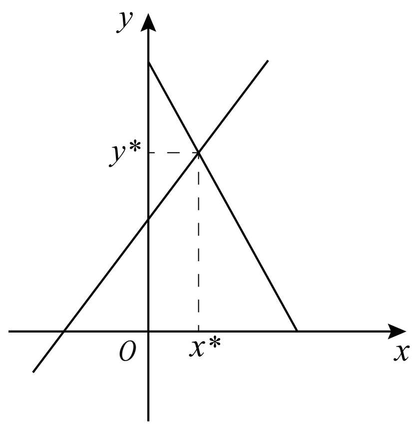

Figure 1.4 Interaction of two straight lines

Algebraically:

Equalizing $y = 60 - 3x$ and $y = 5 + 2x$, we have

$$60 - 3x = 5 + 2x$$

or

$$55 = 5x.$$

Thus, we get the solution for x:

$$x^* = 11.$$

Substituting $x^* = 11$ into either the equation, we have

$$y^* = 60 - 3x^* = 60 - 33 = 27$$

or

$$y^* = 5 + 2x^* = 5 + 22 = 27.$$

1.2.3 Derivatives

The **derivative** of a function f at x, written as $f'(x)$ or $\frac{df}{dx}(x)$, is defined as the **limit** of the difference quotient (if it exists):

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x},$$

and called it is **differentiable** at x.

Thus, ***the derivative of a function at x measures the rate at which the function value changes with respect to a change in the independent variable***, i.e., it is the slope of the tangent to the graph of f at point $(x, f(x))$.

Common rules of derivatives

- sum and difference rule:

$$(f \pm g)'(x) = f'(x) \pm g'(x).$$

- product rule:

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

- quotient rule:

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

- power rule $f(x) = x^k$ (k is real number):

$$f'(x) = kx^{k-1}.$$

- natural logarithmic rule $f(x) = \ln x$:

$$f'(x) = \frac{1}{x}.$$

- exponential rule $f(x) = e^x$ ($e \cong 2.71828\cdots$ is the base of natural logarithm):
$$f'(x) = e^x.$$

- linearity rule (k is real number):
$$(kf)'(x) = kf'(x).$$

- chain rule:
$$(g(f))'(x) = g'(f(x))f'(x).$$

1.2.4 Functions of multiple variables

A function of k variables is a **mapping** $f: X_1 \times X_2 \cdots X_k \to \mathcal{R}$, where X_i for $i = 1, 2, \ldots, k$ are subsetsx of $\mathcal{R}$.

The partial derivative of z with respect to x_i, denoted by $\frac{\partial z}{\partial x_i}$, is defined as:

$$\frac{\partial z}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{\Delta z}{\Delta x_i} = \frac{f(x_1, x_2, \cdots, x_{i-1}, x_i + \Delta x_i, x_i, \cdots, x_n) - f(x_1, x_2, \cdots, x_n)}{\Delta x_i}.$$

Computation is similar to compute a derivative. When computing $\frac{\partial z}{\partial x_i}$, we should view all other variables as constant and view x_i as a variable, compute the derivative (with respect to x_i).

We mostly focus on ***functions of two variables***. In a function of two variables $z = f(x, y)$, x and y are independent variables and z is a dependent variable.

Chapter 2

Market Analysis

Demand-supply analysis is a fundamental and powerful tool that can be applied to a wide variety of interesting and important problems. To name just a few:

- Understanding and predicting how changing world economic conditions affect market price and production;
- Evaluating the impact of government price controls, minimum wages, price supports, and production incentives;
- Determining how taxes, subsidies, tariffs, and import quotas affect consumers and producers.

2.1 Demand

Demand [$D(p)$]: A schedule which shows the relationship between the quantity of a good that consumers are **willing** and **able** to buy and the price of the good, **ceteris paribus**.

Here,

- ***willingness*** to purchase reflects tastes (preferences) of consumers;
- ***ability*** to purchase depends upon income;

- ***ceteris paribus***: all other things remaining constant (preferences, income, prices of other goods, environmental conditions, expectations, size of markets).

Thus, the demand curve, labeled D, reveals how the quantity of a good demanded by consumers depends on its price.

Market demand: The sum of single individual demands.

price of calculators (p)	quantity demanded (millions) of calculators per year
\$20	2
\$15	5
\$10	10
\$5	15
\$1	25

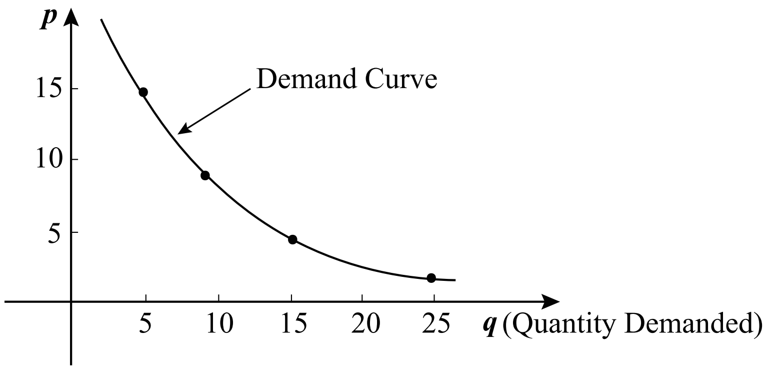

Figure 2.1 The downward-sloping demand curve

Law of Demand: An ***inverse*** relationship between price and quantity demanded. That is, the lower the price, the larger will be the quantity demanded, and vice versa.

A linear demand function can be denoted as $D(p) = ap + b$. An inverse relationship between quantity and price implies $a < 0$.

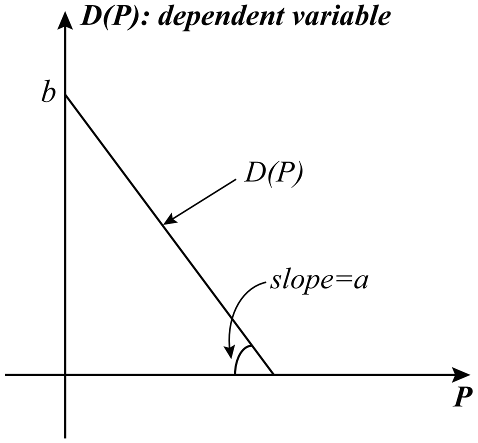

Figure 2.2 A linear demand function

However, ***traditionally, economists reverse the axes when graphing***:

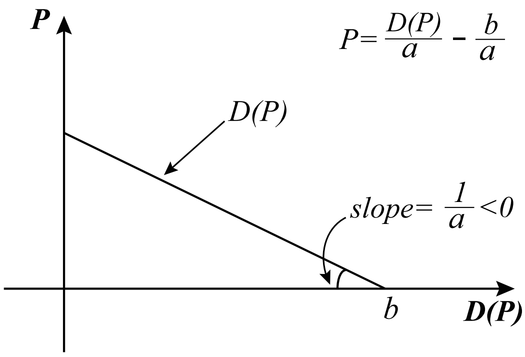

Figure 2.3 A linear demand function with y-axis representing price

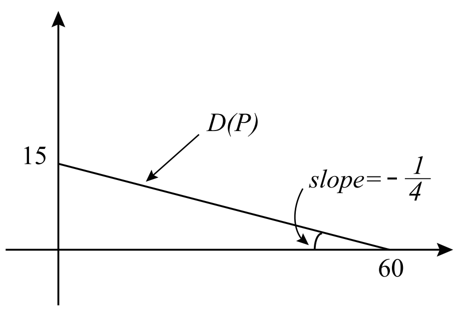

Figure 2.4 A numerical example of linear demand function

Example 2.1 $D(p) = 60 - 4p$. Then $p = -1/4D(p) + 15$ with slope $= -\frac{1}{4}$.

Above, we have imposed the assumption that all other influences on demand are held constant as the price changes. Price change is the only cause of a change in quantity demanded (movement along demand curve).

Changes in Quantity Demanded versus Changes in Demand

(1) **Change in quantity demanded**:

- i) caused by a change in price;
- ii) represented by a movement along the demand curve.

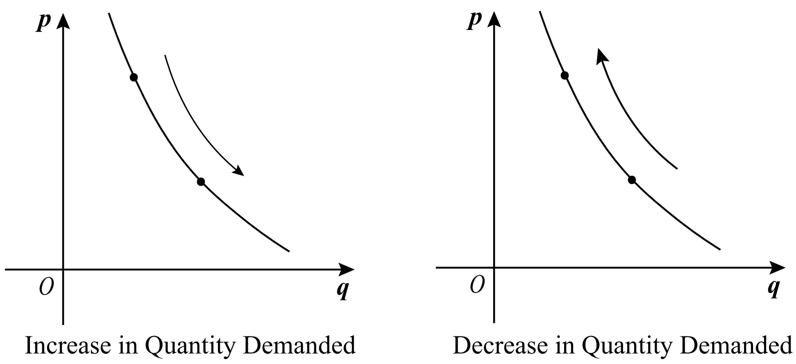

Figure 2.5 Change in quantity demanded: movement along the demand curve

(2) **Change in demand**:

- i) caused by a change in something other than price;
- ii) represented by a shift of the demand curve.

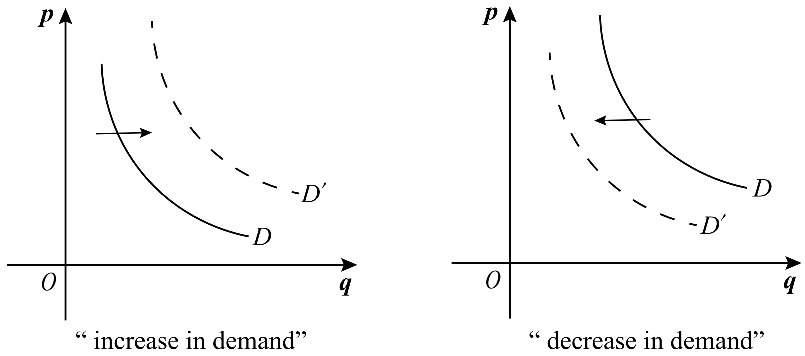

Figure 2.6 Change in demand: the shift of demand curves

Factors causing a change in demand (i.e., factors which shift the demand curve):

a) Size of market as city grows.
E.g., Better marketing ⇒ increase in the number of consumers
⇒ increase in demand as depicted in Figure 2.7.

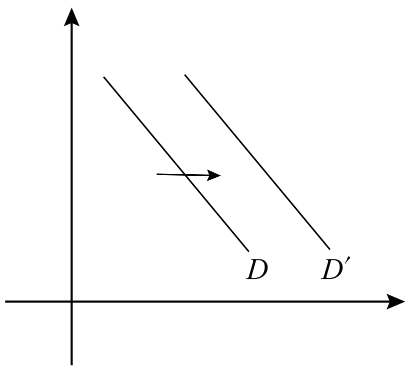

Figure 2.7 Demand curve shifts to the right as size of market grows

b) Income.

- **Normal goods**: As income rises, demand rises; most goods are "normal" goods as also depicted in Figure 2.7.

- **Inferior goods**: As income rises, demand falls; e.g., potatoes, bread (poverty goods) as depicted in Figure 2.8.

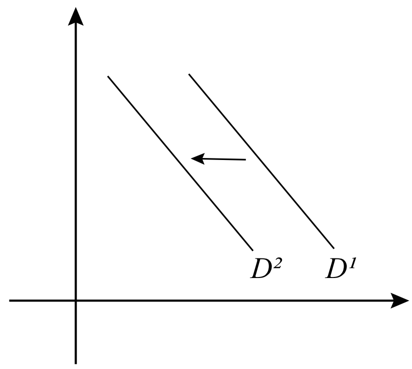

Figure 2.8 Demand curve for inferior goods shifts towards the left as income rises

c) Prices of related goods.

- **substitute goods** (also known as the "**competing goods**"): Two goods for which an increase in the price of one leads to an increase in the demand of the other, e.g., beef and pork as depicted in Figure 2.9.

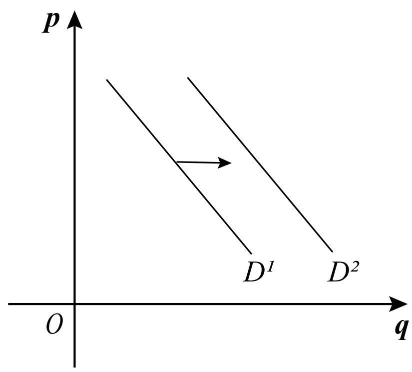

Figure 2.9 Demand curve for pork shifts towards the right as the price of beef increases

— Price of beef rises ⇒ the substitution of pork for beet ⇒ demand for pork rises. Consequently, the *demand curve for pork shifts to the right*.

- **complementary goods**: Two goods for which an increase in the price of one leads to a decrease in the demand of the other, e.g., hamburger and buns as depicted in Figure 2.10.

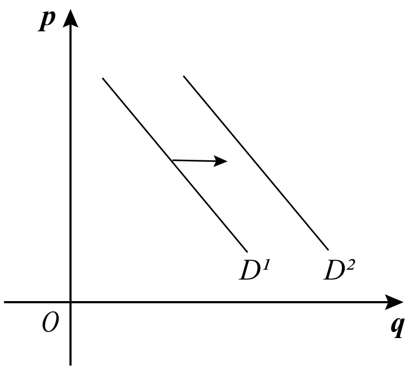

Figure 2.10 Demand curve for buns shifts towards the right as the price of hamburgers increases

—The price of hamburgers falls ⇒ quantity demanded rises ⇒ demand for buns rises.

d) Tastes (preferences)

e.g., cigarette causes cancer ⇒ demand for cigarette falls as depicted in Figure 2.11.

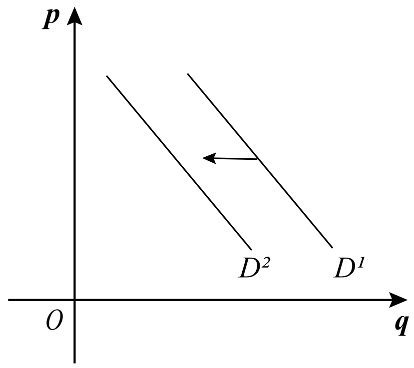

Figure 2.11 Demand curve for cigarette shifts towards the left as people know cigarette causes cancer

e) Expectations

e.g., paper towels go on sale next week ⇒ people will buy them next week ⇒ demand of this week falls as depicted in Figure 2.12.

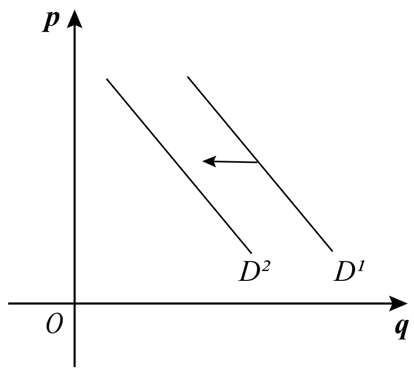

Figure 2.12 Demand curve for paper towels in this week shifts towards the left

f) Environmental conditions

e.g., weather conditions affect demand for air conditioners, ice cream, and winter coats.

2.2 Supply

Supply [$S(p)$]: a schedule which shows the relationship between the quantity of a good that producers are **willing** and **able** to sell and the price of the good, **ceteris paribus**.

The supply curve, labeled S, shows how the quantity of a good offered for sale changes as the price of the good changes. The supply curve is upward sloping: The higher the price, the more firms are able and willing to produce and sell.

Market supply : The sum of supplies of individual firms.

price of calculators	quantity supplied (millions) of calculators per year
\$20	600
\$10	300
\$2	100

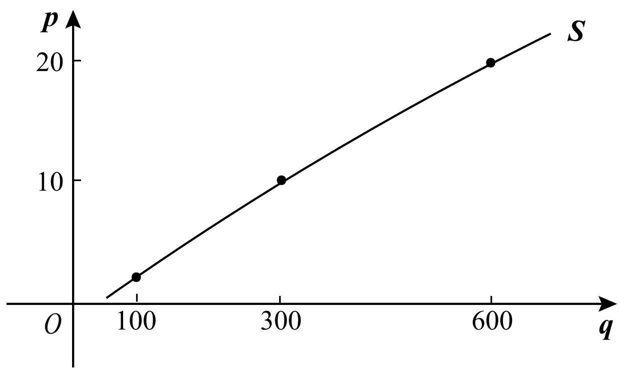

Figure 2.13 A supply curve

Law of Supply: A ***direct*** (i.e., positive) relationship between price and quantity supplied. That is, keeping other factors constant, the higher the price, the larger will be the quantity supplied, and vice versa.

Linear supply function:

$$S(p) = ap + b.$$

Direct relationship $\Rightarrow a > 0$.

Example 2.2 $S(p) = -10 + 6p$. Then, $p = \frac{1}{6}S(p) + \frac{5}{3}$ with slope $= \frac{1}{6}$.

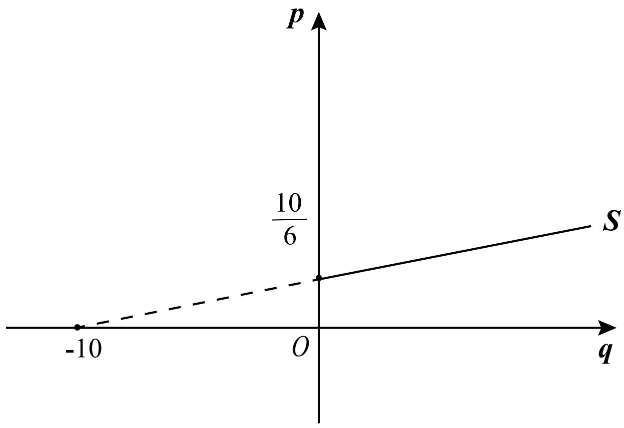

Figure 2.14 A numerical example of supply curve

Why a direct relationship? — ***Substitution of Expansion in Production***:

As the price of a good rises, producers will shift resources into the production of this relatively high priced good and away from production of relatively low priced goods. Alternatively, the producers have incentives to hire extra resources.

Changes in Quantity Supplied versus Changes in Supply

(1) **Change in quantity supplied**:

- i) caused by a change in price;
- ii) represented by a movement along the supply curve.

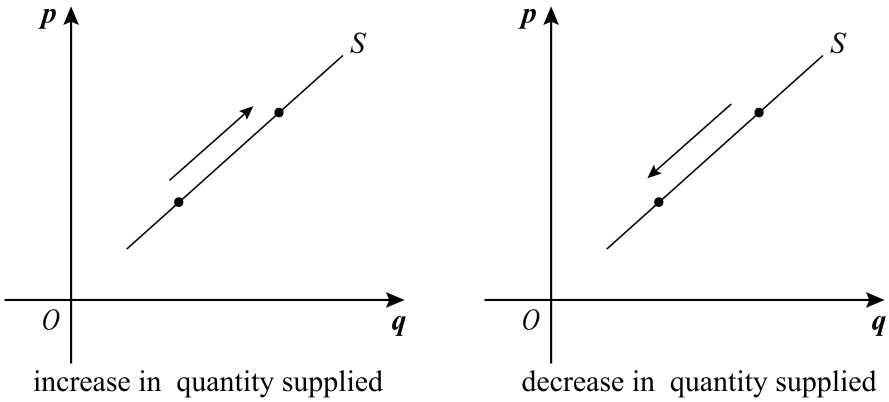

Figure 2.15 Change in quantity supplied: movement along supply curve

(2) **Change in supply**:

- i) caused by a change in something other than price;
- ii) represented by a shift of the supply curve.

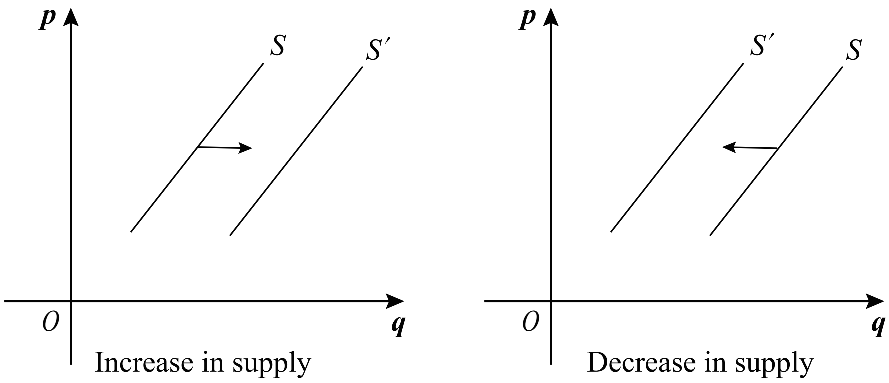

Figure 2.16 Change in supply: shifts in supply curve

Factors Causing a Chance in Supply (factors which shift the supply curve):

a) number of firms
b) prices of related goods

c) technology
d) expectations
e) environmental conditions

Examples of change in supply:

i) prices of resources: increase in wage $\Rightarrow$ increase in costs of production $\Rightarrow$ decrease in supply.
ii) advancement in technology $\Rightarrow$ decrease in costs of production $\Rightarrow$ increase in supply.

2.3 Determination of Equilibrium Price and Quantity

Notations:

- q^d is quantity demanded of commodity;
- q^s is quantity supplied of commodity;
- p is the price of the commodity;

Equilibrium price (also known as the **market clearing price**), denoted by (p^e), is established at the price where quantity demanded equals quantity supplied of the commodity.

Three relevant concepts:

- **Market mechanism**: Tendency in a free market for price to change until the market clears.
- **Surplus**: Situation in which the quantity supplied exceeds the quantity demanded.
- **Shortage**: Situation in which the quantity demanded exceeds the quantity supplied.

When can we use the supply-demand model?

- We are assuming that at any given price, a given quantity will be produced and sold.

- This assumption makes sense only if a market is at least ***roughly competitive***, which means that both sellers and buyers should have ***little market power*** — i.e., little ability ***individually*** to affect the market price.

- Suppose instead that supply were controlled by a single producer — a monopolist. If the demand curve shifts in a particular way, it may be in the monopolist's interest to keep the quantity fixed but change the price, or to keep the price fixed and change the quantity.

Example 2.3 (Example on Calculator (continued)) From Table 2.1, we can know the equilibrium price is $p^e = \$10$. We say that"$p^e$ clears the market".

The equilibrium quantity: $q^e = 3000$ ($q^e = q^s = q^d$).

p	q^d (millions)	q^s (millions)	surplus (+) or shortage (-)
\$20	700	6000	+5300
\$10	3000	3000	0
\$2	6500	1000	-5500

Table 2.1: Market equilibrium of calculators.

The market equilibrium can be also visualized as depicted in Figure 2.17.

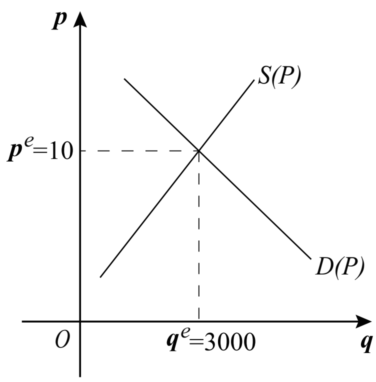

Figure 2.17 The determination of equilibrium price and quantity

Example 2.4 Find the market equilibrium price p^e and equilibrium quantity q^e of the following demand and supply.

$$D(p) = 80 - 4p,$$

$$S(p) = -10 + 6p.$$

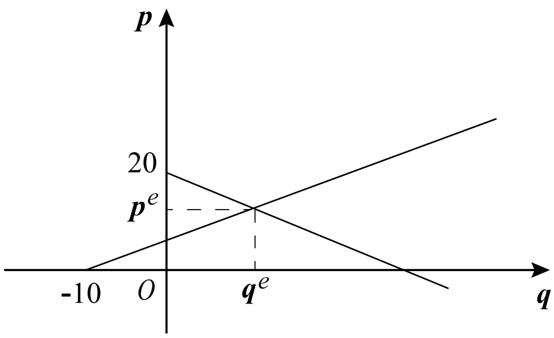

Figure 2.18 Finding market equilibrium: a numerical example

$p = 0$ implies $D(0) = 80$ and $S(0) = -10$. $q^d = 0$ implies $p = 20$, and $q^s = 0$ implies $p = 10/6 = 5/3$. We then can draw the demand curve and supply curve as depicted in Figure 2.18, giving us the market equilibrium price p^e and equilibrium quantity q^e.

Algebraically, equaling $D(p^e) = S(p^e)$, we have

$$80 - 4p^e = -10 + 6p^e,$$

and thus

$$90 = 10p^e,$$

which yields $p^e = 9$.

Substituting $p^e = 9$ into either the equation, we have

$$q^e = D(p^e) = S(p^e) = 80 - 4 \times 9 = 44.$$

2.4 Market Adjustment

The market adjustment mechanism works as follows:

- Suppose $p > p^e$. Then $q^s > q^d$, yielding ***surplus***. Producers compete to extract the surplus by price cutting. Price falls implies that q^s falls and q^d rises. Eventually $q^s = q^d$ at p^e.

- Suppose $p < p^e$. Then $q^d > q^s$, yielding ***shortage***. Consumers compete and force p up, having q^d falls and q^s rises. Eventually $q^d = q^s$ at p^e.

Changes in Supply and/or Demand: Effect on Equilibrium

Example 2.5 (a) Demand increases from D_1 to D_2.

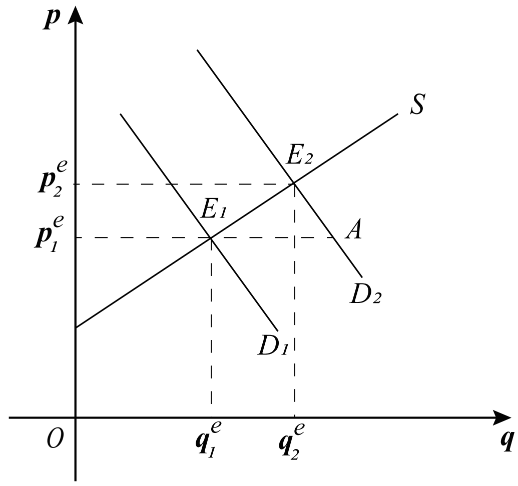

Figure 2.19 The effect of an increase in demand on market equilibrium: resulting in increase in both p^e and q^e

When demand increases (say, due to an increase in income) from D_1 to D_2, we have $q^d > q^s$ at original equilibrium price p_1^e, resulting in shortage, which in turn results in price rises. Thus consumers move from A to E_2, and producers move from E_1 to E_2 as depicted in Figure 2.19.

Result: Demand rises implies that both p^e and q^e rise.

Example 2.6 (b) Supply increases from S_1 to S_2.

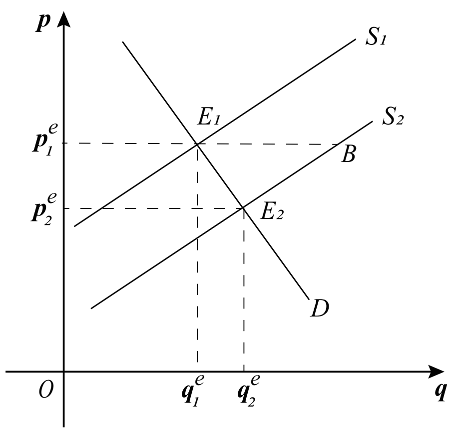

Figure 2.20 The effect of an increase in supply on market equilibrium: resulting in decrease in p^e and increase in q^e

When supply increases (say, due to an increase in productivity or a decrease in cost of production) from S_1 to S_2, we have $q^s > q^d$ at p_1^e, resulting in surplus, which in turn results in price falls. Thus consumers move from E_1 to E_2, and producers move from B to E_2 as depicted in Figure 2.20.

Result:Increase in $S \Rightarrow p^e$ falls and q^e rises.

Example 2.7 (c) Supply and demand rise

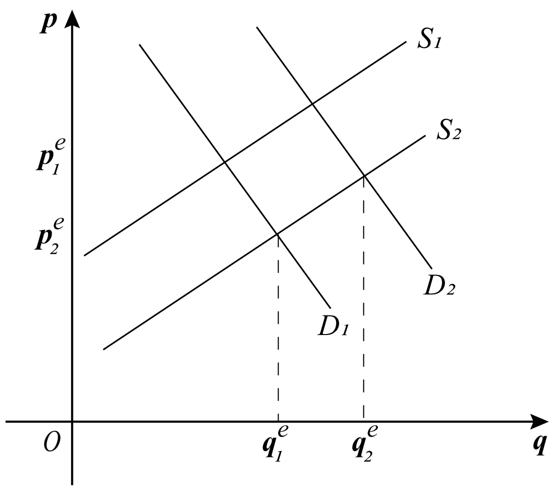

Figure 2.21 The effect of simultaneous increase in demand and supply on market equilibrium: resulting in increase in q^e

Result: Increase in S and D will cause the equilibrium q^e rises as depicted in Figure 2.21, but the equilibrium p^e is uncertain without further information on demand and supply.

Figure 2.22 graphically shows 9 possibilities of the effects of simultaneous changes in demand and/or supply on market equilibrium price and quantity.

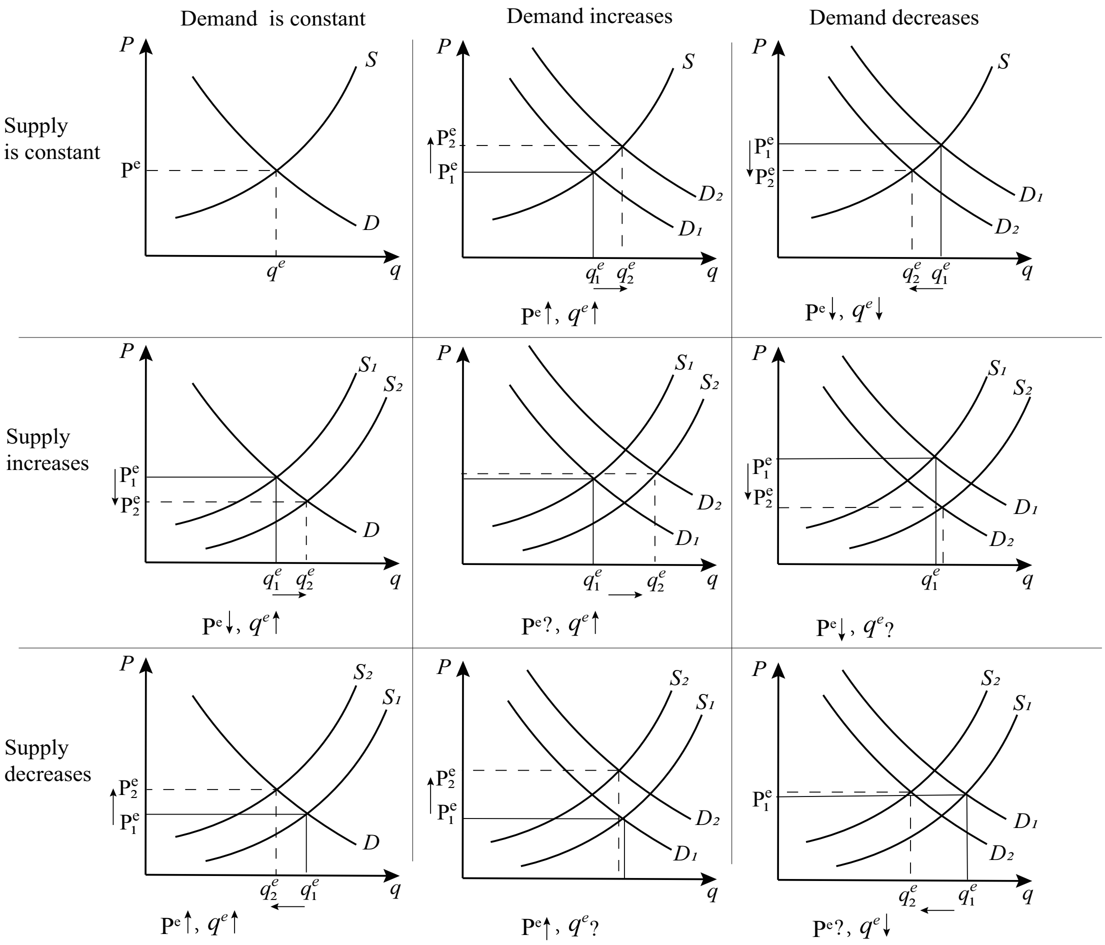

Figure 2.22 Nine possibilities of simultaneous change in demand and supply

Note that as long as demand and supply both change (there are four cases in Figure 2.22), either the equilibrium p^e or quantity q^e is uncertain without further information on demand and supply.

However, with specific demand and supply, we know the changes in equilibrium price and quantity as shown in the following example.

Example 2.8 (Markets for eggs and college (continued)) From 1970 to 2010, the real price of eggs fell by 55 percent, while the real price of college education rose by 82 percent.

The mechanization of poultry farms sharply reduced the cost of producing eggs, shifting the supply curve to the right. The demand curve for eggs shifted to the left as a more health-conscious population tended to avoid eggs. As a result, the real price of eggs fell sharply and egg consumption rose as depicted in (a) Figure 2.23.

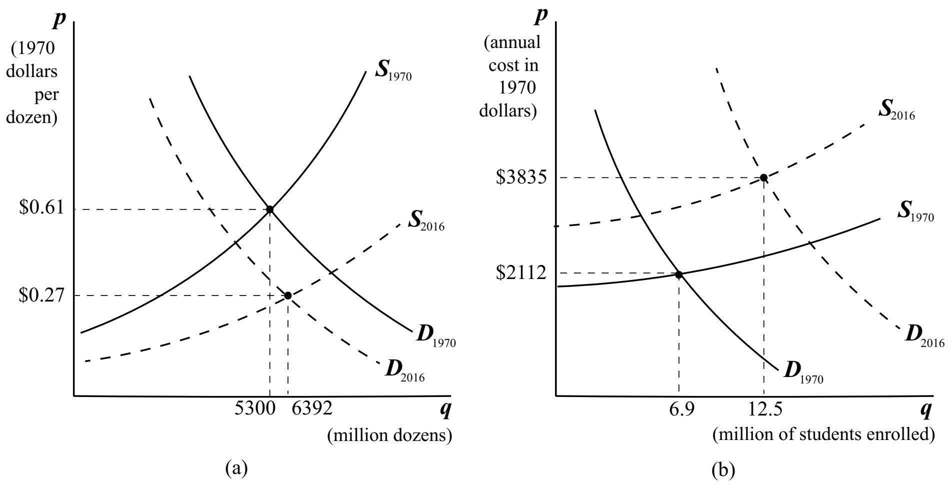

Figure 2.23 (a) Markets for eggs; (b) Markets for college

As for college, increases in the costs of equipping and maintaining modern classrooms, laboratories, and libraries, along with increases in faculty salaries, pushed the supply curve up. The demand curve shifted to the right as a larger percentage of a growing number of high school graduates decided that a college education was essential. As a result, both price and enrollments rose sharply as depicted in (b) Figure 2.23.

Example 2.9 (Explaining wage inequality in the United States) Over the past four decades, the wages of skilled high-income workers have grown substantially, while the wages of unskilled low-income workers have fallen slightly.

From 1978 to 2009, people in the top 20 percent of the income distribution experienced an increase in their average real (inflation-adjusted) pretax household income of 45 percent, while those in the bottom 20 percent saw their average real pretax income increase by only 4 percent.

More surprisingly, by the end of 2021, the top 1% of Americans owned a record-high 32.3% of the country's wealth while the share of wealth held by the bottom 90% of Americans is only 30.2% (so the top 10% of Americans owned almost 70% of the country's total wealth). The top 20% of Americans owned 86% of the country's wealth while the bottom 80% of the population owned 14% in 2020.

Why? One of reasons is the differences in human capital and new technological revolution (such as artificial intelligence, digital economy, etc.). While the supply of unskilled workers (with limited educations or outdated skills) has grown substantially, the demand for them has risen only slightly.

On the other hand, while the supply of skilled workers — e.g., engineers, scientists, managers, and economists — has grown slowly, the demand has risen dramatically, pushing their wages up significantly.

Example 2.10 (Supply and Demand for City Office Space) Since COVID-19, the supply for city office decreases slightly, but the demand decreases substantially (people work from home), so that the average rental price for office fell.

2.5 Elasticity of Demand and Supply

The concept of elasticity has immense importance in consumer's consumption decisions, businessmen's pricing strategies, government policies (such as taxation), and international trade (such as the term of trade between t-wo countries).

Elasticity measures responsiveness of percentage change in quantity demanded (or supplied) to a percentage change in a variable (such as own price, prices of other goods, and income).

Can we use slope of the curve to measure the elasticity? No, because a change in scale can make a curve look flatter or steeper without altering absolute responsiveness. Thus, changing units (say from a dollar to a cent) will alter the slope. For this reason, economists use a unitless measure to measure the elasticity.

2.5.1 Elasticity of Demand

Price elasticity of demand ("own" Price Elasticity): Percentage change in quantity demanded resulting from a percentage change in its price:

$$\begin{aligned} E_d^p &= \frac{\text{percentage change in quantity demanded}}{\text{percentage change in its price}} \\ &= \frac{\Delta Q/Q}{\Delta P/P} \\ &= \frac{P}{Q} \times \frac{\Delta Q}{\Delta p} = \frac{P}{Q} \times \text{the slope of } D, \end{aligned}$$

where "Δ" = "change".

Note that, since demand curves in general slope downward, E_d^p will be negative.

Range of Value for E_d^p:

(a) **Inelastic demand**: $|E_d^p| < 1$.

$$|\frac{\Delta Q/Q}{\Delta P/P}| < 1 \Rightarrow |\Delta Q/Q| < |\Delta P/P|.$$

When demand is inelastic, the quantity demanded is relatively unresponsive to changes in price; e.g., cigarettes.

(b) **Elastic demand**: $|E_d^p| > 1$. That is $|\Delta Q/Q| > ||\Delta P/P|$. When demand is elastic, the quantity demanded is relatively responsive to changes in price; e.g., automobiles in the short run, competing goods.

(c) **Unit elasticity of demand**: $|E_d^p| = 1$, i.e., $|\Delta Q/Q| = |\Delta P/P|$.

(d) **Iso-elastic demand curve Demand curve**: the price elasticity is constant.

(e) **Perfectly inelastic demand**: $|\Delta Q/Q| = 0$ for all changes in price —i.e., totally unresponsive.

(f) **Perfectly elastic demand**: $|E_d^p| = \infty$. A very small percentage change in price leads to a tremendous percentage change in quantity demanded.

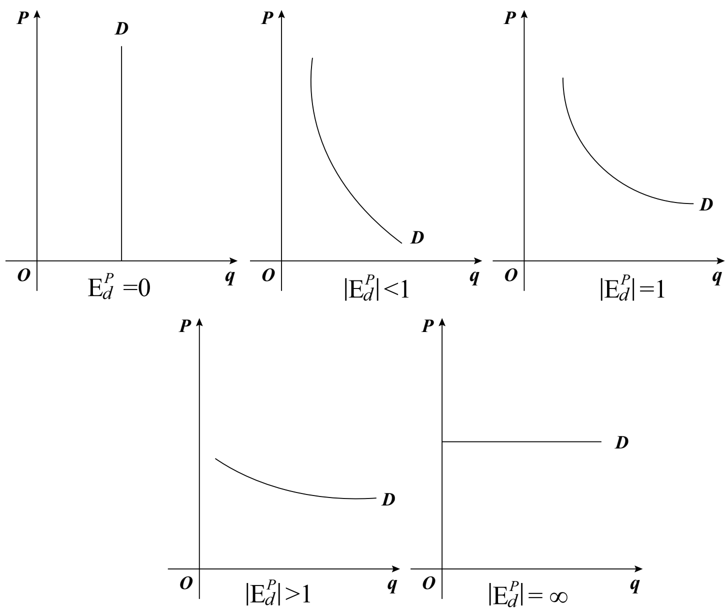

Figure 2.24 Types of demand elasticity

Example 2.11 We first suppose that the price p_x of good x increases from \$1 to \$2 and the quantity demanded q_x changes from 10 to 5.

P	Q
\$1	10
\$2	5

We then have $\Delta P/P = 100\%$, $\Delta Q/Q = -50\%$, and thus $E_d^p = -\frac{50}{100} =$

$-\frac{1}{2}$.

Now suppose that the price decreases from \$2 to \$1 and the quantity demanded changes from 5 to 10.

P	Q
\$2	5
\$1	10

$\Delta P/P = \frac{1-2}{2} = -50\%$, $\Delta Q/Q = 100\%$, so $E_d^p = -2$.

This demonstrates how E_d^p varies according to the P-Q combination which is used as a reference (original or starting) point. To avoid this problem, we use the **midpoint elasticity**, also known as the **arc elasticity** of demand.

Midpoint elasticity of demand: E_d^p between points (P, Q) and (P', Q') is given by

$$E_d^p = \frac{\Delta Q/\frac{1}{2}(Q+Q')}{\Delta P/\frac{1}{2}(P+P')}.$$

This just uses the midpoints as the reference points.

Example 2.12 (Example 2.11 (continued)) Consider the example again.

P	Q
\$2	5
\$1	10

Note that

$$\Delta Q/\frac{1}{2}(Q+Q') = -\frac{5}{\frac{1}{2}\cdot 15} = -\frac{2}{3}$$

and

$$\Delta P/\frac{1}{2}(P+P') = \frac{1}{\frac{1}{2}\cdot 3} = \frac{2}{3},$$

so we have

$$E_d^p = \frac{2/3}{2/3} = -1.$$

Only very special demand curves have constant elasticity between any two points.

Elasticity along Straight Line Demand Curve

The straight line demand function

$$D = a - bP$$

does not have a constant elasticity throughout. Indeed,

$$\begin{aligned} E_d^p &= \frac{\Delta Q/Q}{\Delta P/P} = \frac{P}{Q}\cdot\frac{\Delta Q}{\Delta P} \\ &= \frac{P}{Q} \times \text{the slope of } D \\ &= a \times \frac{P}{Q} \\ &\neq \text{constant} \end{aligned}$$

because $\frac{P}{Q}$ is not constant even though the slope of a linear demand curve, a, is constant.

Moreover, from Figure 2.25, we have

$$|E_d^p| = |\frac{\Delta Q}{\Delta P}|\cdot\frac{P}{Q} = \frac{LB}{AL}\cdot\frac{OL}{LB} = \frac{OL}{AL}.$$

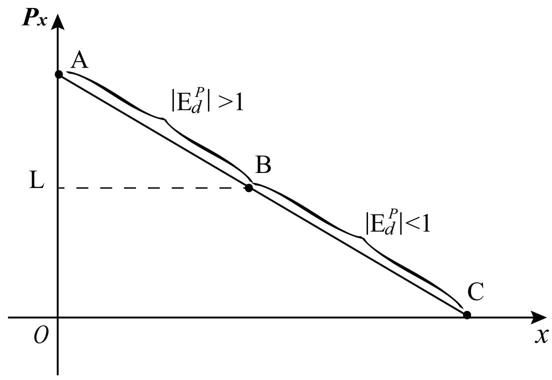

Figure 2.25 Elasticities vary along a linear demand curve

Thus, if B is the midpoint of the demand curve, then $|E_d^p| = 1$. If B is a point above the midpoint, then $AL < OL \Rightarrow |E_d^p| > 1$. If B is a point below the midpoint, $AL > OL \Rightarrow |E_d^p| < 1$.

Therefore, ***along a downward-sloping straight-line demand curve, the price elasticity varies although the slope is constant.***

2.5.2 Elasticity and Total Expenditure

The **total expenditure** (TE) for consuming a good is:

$$TE = P \times Q.$$

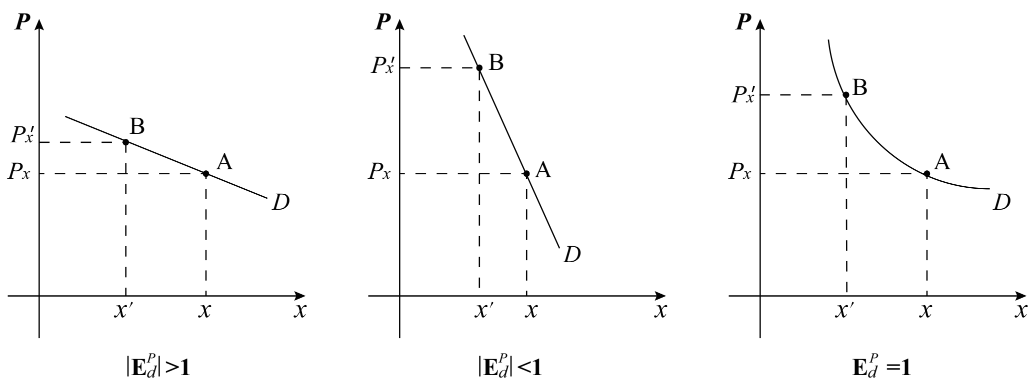

Figure 2.26 Total expenditure under alternative demand elasticities

We are now interested in what happens to the total expenditure $P \times Q$ when P varies. Since P and Q are inversely related by downward-sloping demand, we need further information about the magnitude of the changes in P and Q to determine the effect on $P \times Q$, which is the price elasticity of demand. The relationships are given in Table 2.2.

	$\lvert E_d^p \rvert < 1$	$\lvert E_d^p \rvert > 1$	$\lvert E_d^p \rvert = 1$
P rises	TE rises	TE falls	TE holds constant

Table 2.2: The relationship between elasticity and total expenditure.

In other words, when demand is inelastic, price and total expenditure move in the same direction (this should be easily understand by intuitive such as cigarette). When demand is elastic, price and total expenditure move in opposite directions. When demand is unit elastic, the total expenditure maintains constant when the price varies.

Therefore, when price rises, an inelastic demand implies the total expenditure increases, an elastic demand implies the total expenditure decreases, and consequently, the ***total expenditure reaches its maximum when demand is unit elastic***.

Algebraically, we can verify that these relationships hold true. Suppose $|E_d^p| > 1$ and $P' > P$.

$$
\begin{aligned}
&\frac{\Delta Q/\frac{1}{2}(Q+Q')}{\Delta P/\frac{1}{2}(P+P')} > 1 \\
\Rightarrow &\frac{\Delta Q}{Q+Q'} > \frac{\Delta P}{P+P'} \\
\Rightarrow &(Q-Q')(P+P') > (P'-P)(Q+Q') \\
\Rightarrow &PQ - P'Q' > -PQ + P'Q' \\
\Rightarrow &2PQ > 2P'Q' \\
\Rightarrow &PQ > P'Q'.
\end{aligned}
$$

This shows that if P rises, the total expenditure falls.

Other Elasticities:

- Income elasticity (E_d^I):

$$E_d^I = \frac{\Delta D/D}{\Delta I/I}.$$

- Cross-price elasticity of demand: Let p_y be the price of good y. The cross-price elasticity of demand for good x with respect to the price is defined as

$$E_{d_x}^{p_y} = \frac{\Delta D_x/D_x}{\Delta p_y/p_y}.$$

2.5.3 Price Elasticity of Supply

The Price Elasticity of Supply: a measure of the responsiveness of quantity supplied to a change in price. It is defined as the percentage change in quantity supplied divided by the percentage change in price:

$$E_S^P = \frac{\Delta S/S}{\Delta P/P} = \frac{P}{S} \times \frac{\Delta S}{\Delta P} = \frac{P}{S} \times \text{the slope of } S.$$

We have the following typical cases:

- $E_S^P > 0$ since supply slopes upwards.
- If $E_S^P > 1$, then supply is **elastic**.
- If $E_S^P < 1$, then supply is **inelastic**.
- If $E_S^P = 1$, then supply is **unit elastic**.
- If $E_S^P = 0$, then supply is **perfectly inelastic**.
- If $E_S^P = \infty$, then supply is **perfectly elastic**.

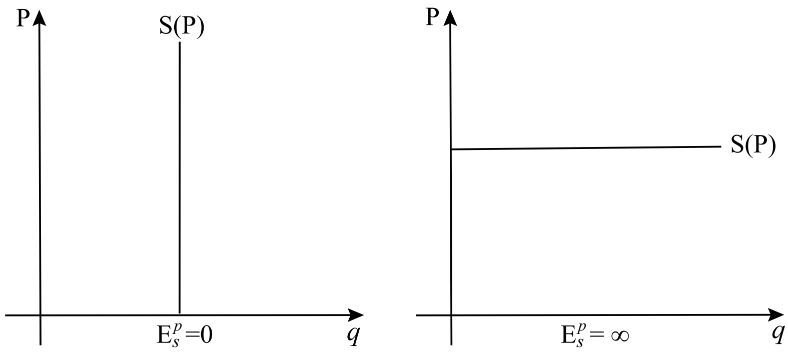

Figure 2.27 Perfectly inelastic and perfectly elastic supplies

Example 2.13 (Market for Wheat) During recent decades, changes in the wheat market had major implications for both American farmers and U.S. agricultural policy.

To understand what happened, let's examine the behavior of supply and demand beginning in 1981.

$$Q_s = 1800 + 240P$$

$$Q_d = 3550 - 266P$$

By setting the quantity supplied equal to the quantity demanded, we can determine the market-clearing price and the equilibrium quantity of wheat for 1981:

$$P = \$3.46,$$

and

$$Q = 2630.$$

We then can use the demand curve to find the price elasticity of demand at market equilibrium:

$$E_D^P = \frac{P}{Q} \times \frac{\Delta Q_D}{\Delta P} = \frac{3.46}{2630}(-266) = -0.35$$

and so demand is inelastic at equilibrium.

We can likewise calculate the price elasticity of supply at market equilibrium

$$E_S^P = \frac{P}{Q} \times \frac{\Delta Q_S}{\Delta P} = \frac{3.46}{2630}(240) = 0.36.$$

Suppose that a drought caused the supply decreases significantly, which pushes the price up to \$4.00 per bushel. In this case, the quantity demanded would fall to 3550 - (266)(4.00) = 2486 million bushels. At this price and quantity, the elasticity of demand would be

$$E_D^P = \frac{P}{Q} \times \frac{\Delta Q_D}{\Delta P} = \frac{4.00}{2486}(-266) = -0.43$$

In 2007, demand and supply were

$$Q_s = 1400 + 115P$$

$$Q_d = 2900 - 125P$$

The market-clearing price and the equilibrium quantity are then

$$P = \$6,$$

and

$$Q = 2150.$$

Dry weather and heavy rains, combined with increased export demand caused the price to rise considerably. You can check to see that, at the 2007 price and quantity, the price elasticity of demand was -0.35 and the price elasticity of supply 0.32. Given these low elasticities, it is not surprising that the price of wheat rose so sharply.

2.5.4 Short-Run versus Long-Run Elasticities

Consumption of Durables and Nondurables

Consumer expenditures include **durable goods** (automobiles, appliances, furniture, etc.) and **nondurable goods** (fuel, food, clothing, services, etc.).

For nondurable goods, the price elasticity of demand is larger in the long run than in the short run. For a durable good, the opposite is true. The short-run price elasticity of demand will be much larger than the long-run elasticity for durable goods. These relationships are also true for income elasticity of demand.

We can verify these assessments by the following example.

Example 2.14 (Elasticities of Demand for Gasoline and Automobiles) Consider the price and income elasticities of demand for gasoline and automobiles, which are given in Table 2.3 and Table 2.4, respectively.

Elasticity for Gasoline	1	2	3	5	10
price	-0.2	-0.3	-0.4	-0.5	-0.8
income	0.2	0.4	0.5	0.6	1.0

Table 2.3: The price and income elasticities of demand for gasoline: Number of years allowed to pass following a price or income change

In the short run, an increase in price of gasoline has only a small effect on the quantity of gasoline demanded. Motorists may drive less, but they will not change the kinds of cars they are driving overnight. In the longer run, however, because they will shift to smaller and more fuel-efficient cars, the effect of the price increase will be larger. Demand for gasoline, therefore, is more elastic in the long run than in the short run.

The opposite is true for automobile demand. If the price of automobile increases, consumers initially defer buying new cars; thus annual quantity demanded for automobile falls sharply. In the longer run, however, old cars wear out and must be replaced; thus annual quantity demanded picks

Elasticity for Automobiles	1	2	3	5	10
price	-1.2	-0.9	-0.8	-0.6	-0.4
income	3.0	2.3	1.9	1.4	1.0

Table 2.4: The price and income elasticities of demand for automobiles: Number of years allowed to pass following a price or income change

up. Demand for automobiles, therefore, is less elastic in the long run than in the short run.

2.6 Effects on Government Intervention: Price Controls

Markets can be thought of as a **self-adjustment mechanism**; they automatically adjust to any change affecting the behavior of buyers and sellers in the market. But for this mechanism to operate effectively, the price must be free to move in response to the interplay of supply and demand. When the government steps in to regulate prices, the market does not function in the same way.

There are two types of price controls: **price ceiling** and **price floor**.

Price Ceiling:

A price above which buying or selling is illegal. It is aimed to help consumers.

Allocation Methods:

i) first come, first served;
ii) rationing (using coupons).

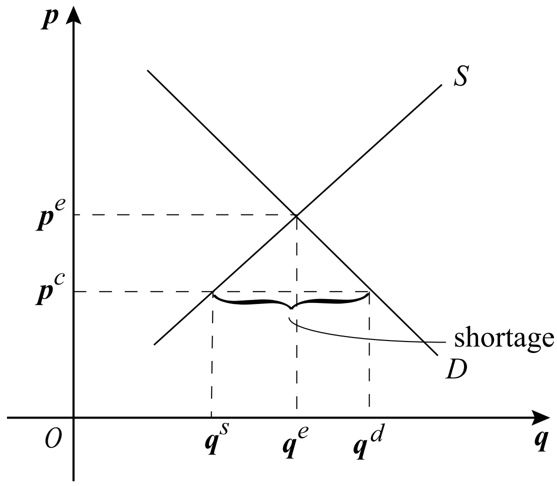

Figure 2.28 Price ceiling: the emergence of shortage

Effects of Price Ceiling:

a. in general it results in shortage;
b. there is a tendency to form a black market;
c. bad service and bad quality of goods;
d. production is reduced;
e. provide wrong information about production and consumption;
f. it hurts producers who provide goods, and some consumers gain from the price ceiling but other may be worse off.

Price Floor (also known as Price Support):

A price below which buying or selling is prohibited. It is aimed to help producers. Examples include setting prices of agricultural products and minimum wage rate.
Effects of Price Floor:

a. in general it results in surplus;
b. provide unnecessary service;

c. lead to over investment;

d. provide wrong information about production and consumption.

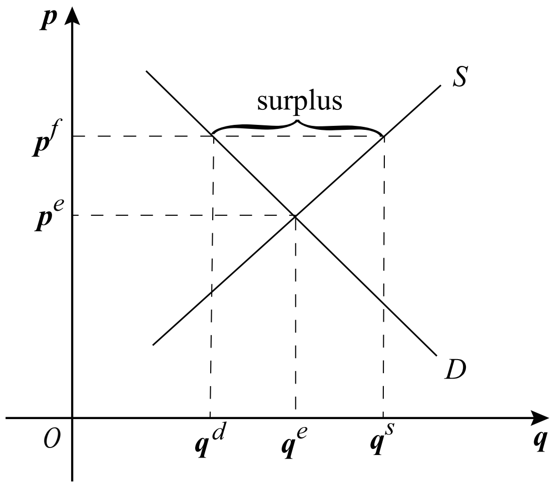

Figure 2.29 Price floor: the emergence of surplus

Methods for Maintaining Price Support:

i) the government purchases surplus, and then the total revenue of the producer $= p^f \times q^s$.

ii) output is restricted at q^d, and then the total revenue of the producer $= p^f \times q^d$.

Besides the effects mentioned for these two types of price controls, we will discuss they also result in welfare losses.

Example 2.15 (Price Control and Wheat) Suppose that the demand and supply of wheat are respectively given by

$$D(p) = 90 - 20p,$$

$$S(p) = -15 + 10p.$$

a) Find the market equilibrium price and quantity.

Setting $D(p) = S(p)$, we have $90 - 20p = -15 + 10p$ which gives us $p^e = 105/30 = 3.5$ and $q^e = 20$.

b) Suppose a price support is set at \$4. What is the surplus?

Since

$$D(4) = 90 - 20 \times 4 = 10,$$

$$S(4) = -15 + 10 \times 4 = 25,$$

so the surplus is given by

$$S(4) - D(4) = 25 - 10 = 15.$$

Part II

Demand Side of Market

Part 2 presents theoretical core of consumer behavior and individual and market demands.

Chapter 3

Theory of Consumer Choice

This chapter discusses the theory of consumer choice – a bedrock foundation of economics. It can be viewed as a typical situation of how an individual makes an independent decision. In practice, corporate decision and public policy require an understanding of the theory of consumer behavior: the explanation of how consumers allocate incomes to the purchase of different goods.

A consumer can have numerous characteristics, such as gender, appearance, age, lifestyle, wealth, preferences, ability, etc. Which of the above are critical in determining the consumer's optimal choice? In principle, a consumer's choice is determined by the consumer's **subjective preference** subject to **objective restrictions**, typically, budget constraints.

3.1 Budget Constraints

Budget constraints: Constraints that consumers face as a result of limited incomes.

3.1.1 Budget Line

The **budget line**: A straight line representing all possible combinations of goods that a consumer can obtain at given prices by spending a given income, namely, the total expenditure of consumptions is equal to income.

A budget line with two commodities, as depicted in Figure 3.1, is:

$$p_x x + p_y y = I, \tag{3.1}$$

where p_x and p_y represent prices of goods x and y, and hence $p_x x + p_y y$ stands for total expenditure which is equal to income I.

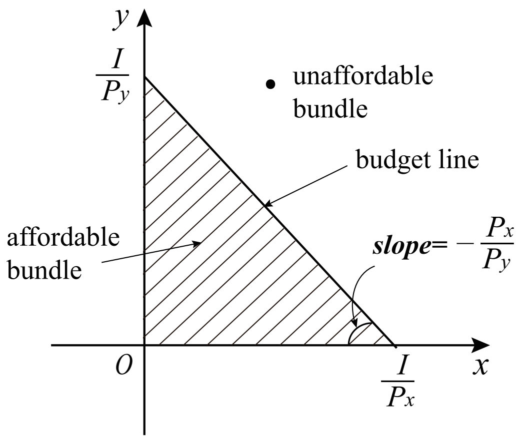

Figure 3.1 Budget constrains and budget line

We can rewrite the budget line (3.1) as

$$y = \frac{I}{p_y} - \frac{p_x}{p_y} x, \tag{3.2}$$

where the slope of (3.1) is $-\frac{p_x}{p_y}$, x-axis intercept is $\frac{I}{p_x}$ representing the maximum units of good x that can be purchased, and y-axis intercept is $\frac{I}{p_y}$ representing the maximum units of good y that can be purchased.

Example 3.1 Two goods x and y, with prices $p_x = 10$ and $p_y = 5$. Income is $I = 100$. So the budget line is

$$10x + 5y = 100.$$

The slope of budget line is $-\frac{p_x}{p_y} = -10/5 = -2$, y-axis intercept is $\frac{I}{p_y} = \frac{I}{p_y} = 100/5 = 20$, and x-axis intercept is $\frac{I}{p_x} = 100/10 = 10$.

Combination:	$p_x \times$ unit of x	+	$p_y \times$ unit of y	=	income
a.	10×10	+	5×0	=	$100
b.	10×8	+	5×4	=	$100
c.	10×6	+	5×8	=	$100
d.	10×4	+	5×12	=	$100
e.	10×2	+	5×16	=	$100
b.	10×0	+	5×20	=	$100

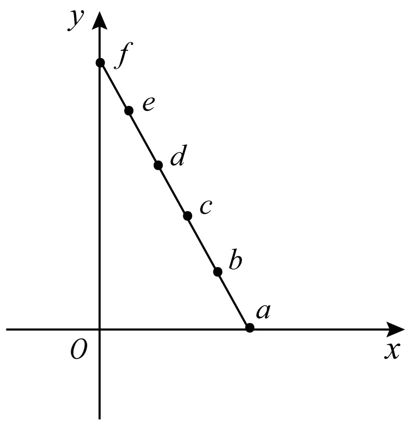

Figure 3.2 A budget line with two consumption goods

3.1.2 Changes in Budget Line

What if prices and income increase at the same rate? If we double both prices and income, then $2p_x x + 2p_y y = 2I$, and nothing changes.

What if only one price changes? If p_y in the above example is changed to $p_y^* = 10$, and p_x and I are the same as before, the new budget line is $10x + 10y = 100$, with slope $-\frac{p_x}{p_y} = -1$. That is, a change in the price of one good causes the budget line to rotate about one intercept.

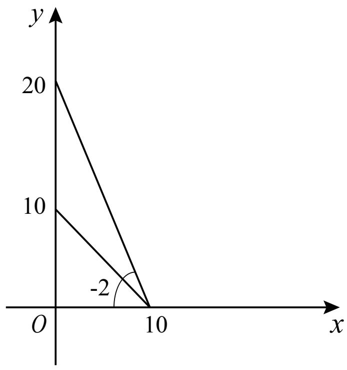

Figure 3.3 The effect of a change in price on budget line

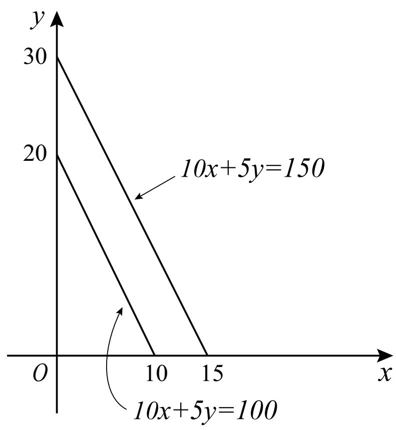

Figure 3.4 The effect of a change in income on budget line

What if only income changes? If the income in the above example changes from \$100 to \$150, the new budget line is $10x + 5y = 150$. Then a change in income (with prices unchanged) causes the budget line to shift parallel to the original line.

3.2 Consumer Preferences

The consumer is assumed to have preferences over bundles (baskets) of goods. Suppose that there are 2 goods available, x and y, and two bundles of these goods, A and B, where each bundle contains a given amount of x and y:

$$A = (x^A, y^A),\ B = (x^B, y^B).$$

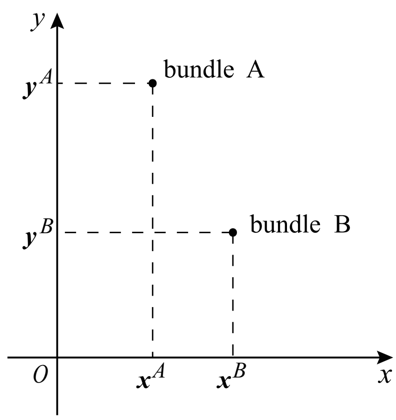

Figure 3.5 Two alternative consumption bundles A and B

3.2.1 Basic Assumptions on Preferences

We make the following assumptions on the consumer's preferences.

i) **Completeness**: consumers can compare and rank all possible baskets. Between any two bundles, the consumer can only make one of the following statements:

A is preferred to B ($A \succ B$),

B is preferred to A ($B \succ A$), or

A is indifferent to B ($A \sim B$).

By ***indifferent*** we mean that the consumer will be equally satisfied with either basket. Note that these preferences ignore costs. A consumer might prefer steak to hamburger but buy hamburger because it is cheaper.

ii) **Transitivity**:

$A \succ B$ and $B \succ C$ imply $A \succ C$.

Transitivity is normally regarded as necessary for consumer consistency.

iii) **More is preferred to less**: goods are desirable;

if $A = (x^A, y^A)$, $B = (x^A, y^A + c)$, $c > 0$, then $B \succ A$.

Consequently, consumers always prefer more of any good to less. In addition, consumers are never satisfied or satiated; ***more is always better, even if just a little better***.

Indifference curve: a curve representing all combinations of market baskets that provide a consumer with the same level of satisfaction. Thus, a consumer is indifferent between any two bundles that lie on the same indifference curve.

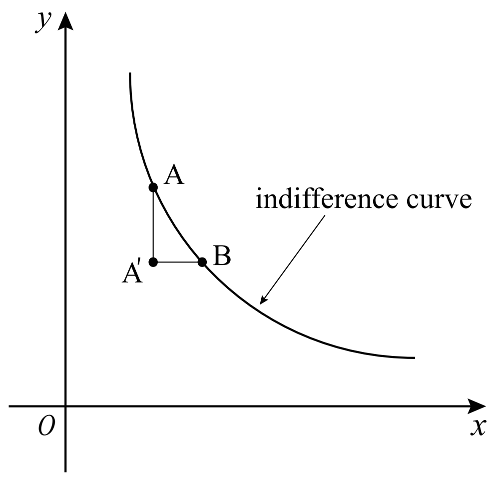

Figure 3.6 The process of formulating an indifference curve

To find an indifference curve, we can start from bundle A. Subtracting one unit of good y, puts us at a point A', strictly less preferred to A by assumption iii), i.e., $A' \prec A$. However, if we add more x to A', we know that

A' will be less preferred to the resulting new point. If we add "enough" x to A', then we find a point B such that $A \sim B$, as shown in Figure 3.6, etc.

Similarly, we can trace out a family of indifference curves, namely, an **indifference map** — i.e., a graph containing a set of indifference curves showing the market baskets among which a consumer is indifferent. In Figure 3.7, we have $A \sim B$, $\bar{A} \sim \bar{B}$, $\bar{\bar{A}} \sim \bar{\bar{B}}$, but any point on U^1 is preferred to any point on U^0, and is less preferred to any point on U^2 by assumptions ii) and iii).

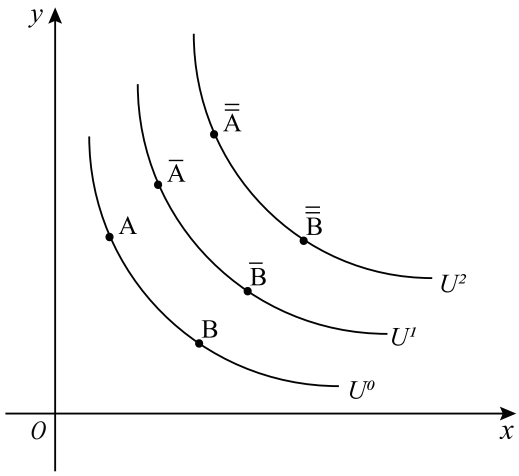

Figure 3.7 Indifference map: a set of indifference curves

3.2.2 Properties of Indifference Curves

There are two crucial properties of indifference curves:

i) ***Indifference curves slope downward*** by assumption ii).
ii) ***Indifference curves cannot intersect***.

Suppose they did intersect, note that $A \sim B$, $C \sim B$, $C \succ A \Rightarrow C \succ B$ but actually $C \sim B$, hence a contradiction.

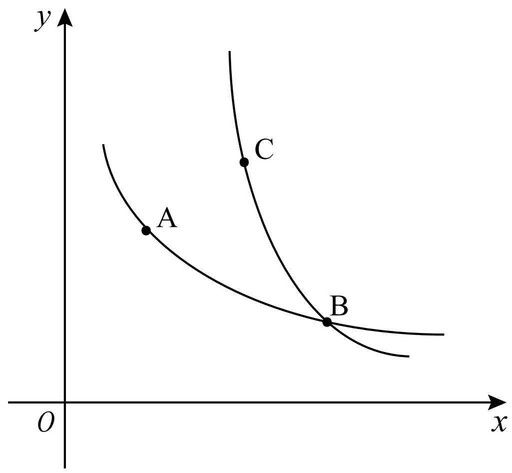

Figure 3.8 Indifference curves cannot intersect

3.2.3 Marginal Rate of Substitution

The slope of an indifference curve reveals the so-called marginal rate of substitution of one good for another good.

The **marginal rate of substitution of** x **for** y (MRS_{xy}): Represent the maximum of units of y that must be given up for one extra unit of x if the consumer is to remain indifferent. In short, it is ***the rate at which one can be substituted for another with satisfaction remaining constant.***

Formally, $MRS_{xy} = \left|\frac{\Delta y}{\Delta x}\right|$ *is the absolute value of the slope of the indifference curve*, where"Δ" stands for small changes. That is, the magnitude of the slope of an indifference curve measures the consumer's marginal rate of substitution (MRS) between two goods x and y.

Indifference curves generally satisfy the following important property.

Diminishing MRS: The amount of good y that the consumer is willing to give up for one additional unit of x decreases as units of x obtained increase.

Diminishing MRS implies that the MRS falls as we move down the indifference curve, which is equivalent to the strict convexity of consumer preferences:

Strict Convexity: Indifference curves are strictly convex to the origin, resulting in diminishing MRS, as depicted in Figure 3.9.

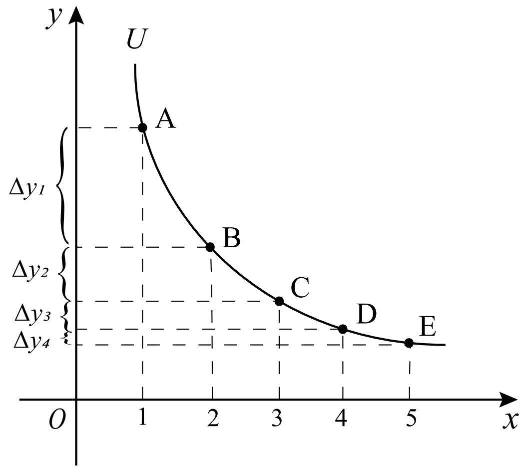

Figure 3.9 The shape of indifference curves: convex to the origin

In Figure 3.9, as at A, y is abundant, x is scarce; thus, the consumer is willing to give up a relatively large amount of the plentiful good to obtain the scarce. At E, y is scarce, x is abundant; the consumer is willing to give up less amount of y for another unit of x.

Example 3.2 Consider the numerical example depicted in Figure 3.10.

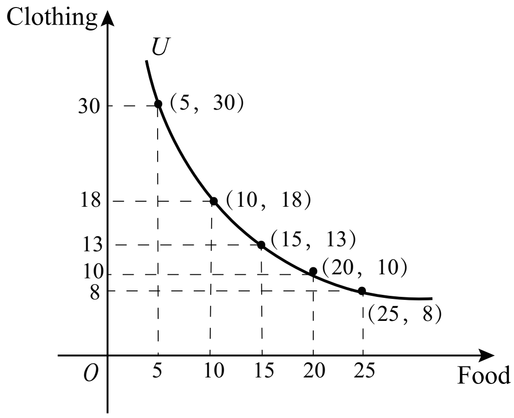

Figure 3.10 MRS between clothing and food

Clothing (C)	Food (F)	MRS of F for C	
30	5		
18	10	$\lvert\frac{\Delta C}{\Delta F}\rvert = \frac{30-18}{10-5} = 2.4$	$\downarrow$
13	15	$\lvert\frac{\Delta C}{\Delta F}\rvert = \frac{18-13}{15-10} = 1.0$	$\downarrow$
10	20	$\lvert\frac{\Delta C}{\Delta F}\rvert = \frac{13-10}{20-15} = 0.6$	$\downarrow$
8	25	$\lvert\frac{\Delta C}{\Delta F}\rvert = \frac{10-8}{25-20} = 0.4$	$\downarrow$
7	30	$\lvert\frac{\Delta C}{\Delta F}\rvert = \frac{8-7}{30-25} = 0.2$	$\downarrow$

The convexity of preferences implies that individuals want to diversify their consumptions (the consumer prefers averages to extremes), and thus, convexity can be viewed as the formal expression of basic measure of ***economic markets for diversification***.

3.2.4 Special Shapes of Indifference Curves:

An indifference curve *may not be strictly convex* to the origin (although it is convex) for some special shapes of indifference curves.

i) **Perfect substitutes** (linear indifference curves): Since the slope of a straight line is constant, MRS is constant.

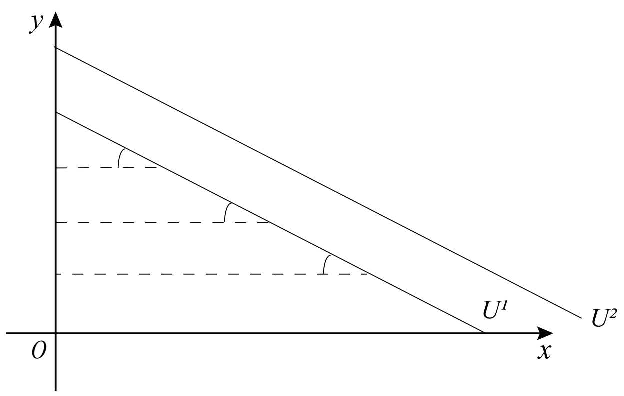

Figure 3.11 Perfect substitutes: constant MRS

ii) **Perfect complements**: two goods for which the MRS is zero or infinite; the indifference curves are shaped as right angles, e.g., x = left shoe,

$y =$ right shoe.

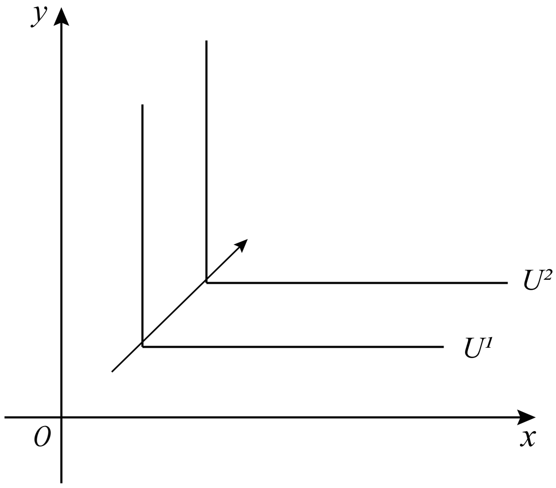

Figure 3.12 Perfect complements: shaped as right angles

iii) Only the consumption of good y matters; see Figure 3.13.

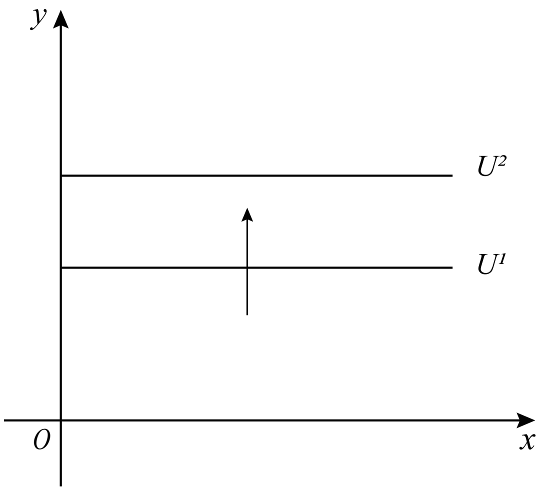

Figure 3.13 Horizontal indifference curves

iv) Only the consumption of good x matters; see Figure 3.14.

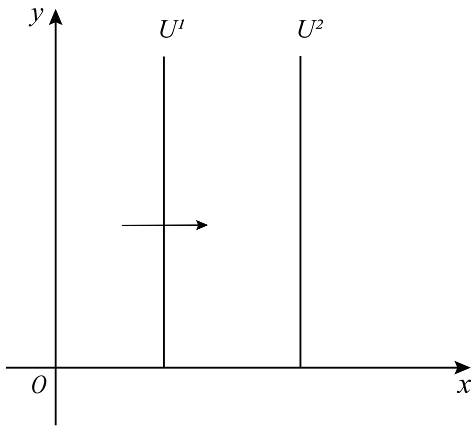

Figure 3.14 Vertical indifference curves

v) Bads: Bad good for which less is preferred rather than more; see Figure 3.15.

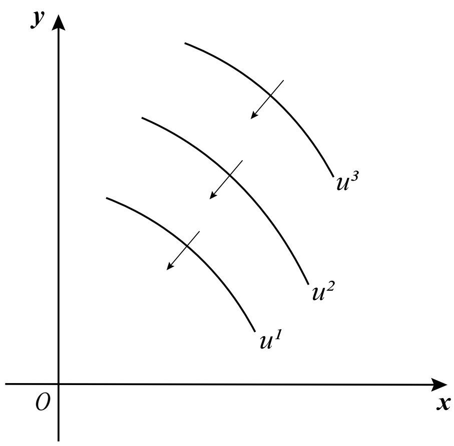

Figure 3.15 Vertical indifference curves

3.2.5 Utility Functions

Indifference curves, which represent consumer preferences, enable us to rank commodity bundles. That is, as shown in Figure 3.16, $A \prec B$, and

$B \prec C$. Sometimes it is convenient to summarize these rankings with a numerical score. Therefore, #1 $= C$, #2 $= B$, #3 $= A$.

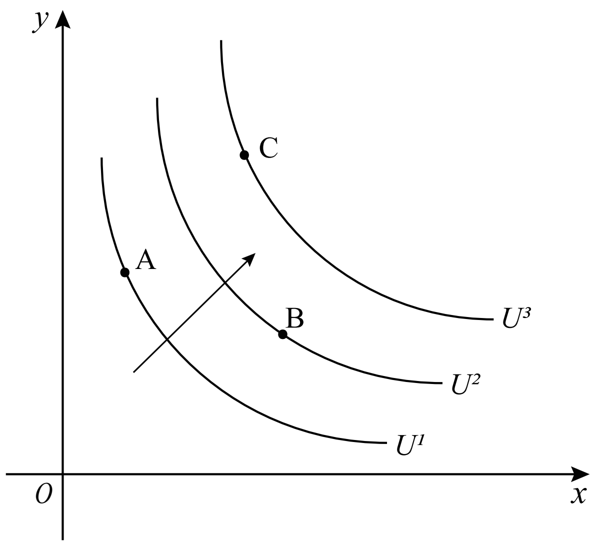

Figure 3.16 Ranking alternative market baskets using indifference curves

A **utility function** can be derived by a set of indifference curves, each with a numerical indicator, which represents the consumer's preferences. That is, $U(x, y)$ assigns a number to commodity bundle (x, y) such that whenever

(a) $U(x, y) > U(\bar{x}, \bar{y})$, then bundle (x, y) is preferred to bundle $(\bar{x}, \bar{y})$.

(b) $U(x, y) = U(\bar{x}, \bar{y})$, then the consumer is indifferent between (x, y) and $(\bar{x}, \bar{y})$.

Thus $U(x, y)$ can be used to rank commodity bundles, in which situation it is known as the **ordinal utility function**. If it is used to measure the levels of utility, it is known as the **cardinal utility function**.

- Ordinal utility function: Utility function that generates a ranking of market baskets in order of most to least preferred.

- Cardinal utility function: Utility function describing by how much one market basket is preferred to another.

The notion of cardinal utility function can be used to measure the happiness of people.

Example 3.3 (Can Money Buy Happiness?) A cross-country comparison shows that individuals living in countries with higher GDP per capita are on average happier than those living in countries with lower per-capita GDP.

The ordinal feature of utility function implies that the *utility function is unique up to arbitrarily monotonic transformations*. In other words, for any strictly increasing function f (such as the squared function) and the utility function U, its composition of $f(U)$ (so that $f = U^2$) represents the same preference ordering (indifference map) as U.

To find an indifferent curve from a utility function, we provide the following examples:

Example 3.4 Suppose

$$U(x, y) = x + 2y.$$

Along an indifferent curve the consumer is indifferent between alternative bundles. That is, along an indifference curve $U(x, y) = \bar{U}$ is constant.

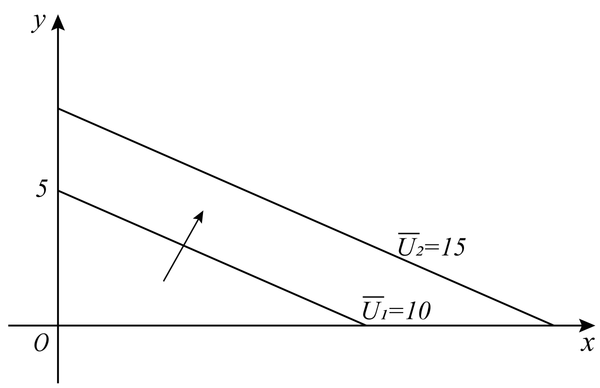

Figure 3.17 Tracing linear indifference curves

Suppose $\bar{U_1} = 10$. To trace an indifference curve, we find all combinations of (x, y) which satisfy $U(x, y) = \bar{U_1} = 10 \Rightarrow x + 2y = 10$.

Suppose $\bar{U_2} = 15 \Rightarrow x + 2y = 15$.

Since $U(x, y) = 10 < U(x, y) = 15$, the latter indifference curve represents bundles preferred to the original set. That is consistent with the indifference curve being further from the origin.

Example 3.5 Suppose

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

At $\bar{U} = 50 \Rightarrow xy = 50$. Bundles which satisfy this equation include

- $x = 25$ and $y = 2$,
- $x = 10$ and $y = 5$,
- $x = 5$ and $y = 10$,
- $x = 2$ and $y = 25$.

We then have the indifference curve as depicted as in Figure 3.18.

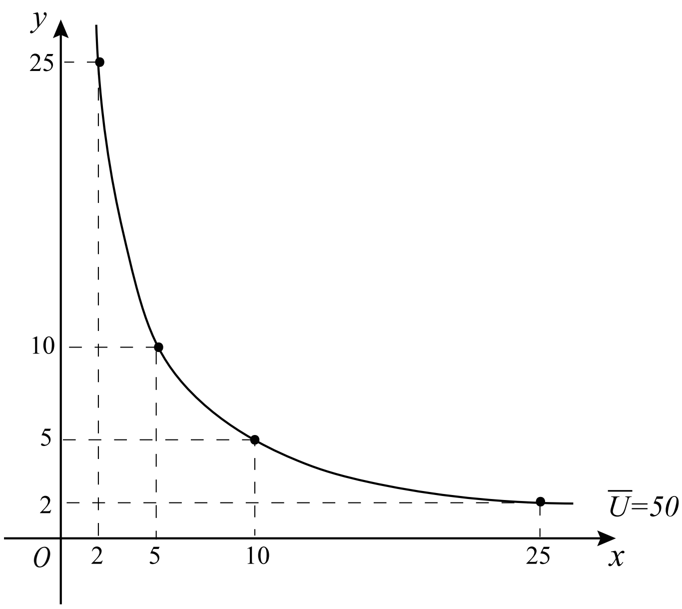

Figure 3.18 Tracing convex indifference curves

Computing MRS from Utility Function

If we know a particular utility function, we can compute MRS of the corresponding indifference curve. Indeed, along an indifference curve represented by a utility level $\bar{U}$, we have $\Delta\bar{U} = 0$, hence

$$0 = MU_x \times \Delta x + MU_y \times \Delta y.$$

We then have

$$\frac{\Delta y}{\Delta x} = -\frac{MU_x}{MU_y}.$$

Since the slope of an indifference curve is $\frac{\Delta y}{\Delta x} = -MRS_{x,y}$, we have

$$MRS_{x,y} = \frac{MU_x}{MU_y}.$$

Example 3.6 (Example 3.4 continued) Suppose

$$U(x, y) = x + 2y.$$

Since $MU_x = 1$ and $MU_y = 2$, we have

$$MRS_{xy} = \frac{MU_x}{MU_y} = \frac{1}{2}.$$

Example 3.7 (Example 3.5 continued)

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

Since $MU_x = y$ and $MU_y = x$, we have

$$MRS_{xy} = \frac{MU_x}{MU_y} = \frac{y}{x}.$$

Now consider the squared transformation of $U(x, y) = xy$ so that

$$U(x, y) = x^2y^2$$

Since $MU_x = 2xy^2$ and $MU_y = 2x^2y$, we have

$$MRS_{xy} = \frac{MU_x}{MU_y} = \frac{2xy^2}{2x^2y} = \frac{y}{x},$$

which is the same as the one for $U(x, y) = xy$. Once again, it shows that the invariance of utility function to a monotonic transformation.

3.3 Consumer's Optimal Choice

The consumer chooses a bundle (x, y) to maximize her preference subjec-t to the budget constraint. That is, the maximizing market basket must satisfy two conditions by Assumptions 1-3:

1. It must be located on the budget line.
2. It must give the consumer the most preferred combination of goods and services.

There are two cases to consider.

Case 1: Interior solution. Indifference curves are strictly convex and do not cross the axes.

From Figure 3.19, one can see

- C - too expensive and hence is not affordable for the consumer,
- $B \succ A$,
- hence B is the utility-maximizing bundle; at this point, the budget line and indifference curve U_2 are tangent.

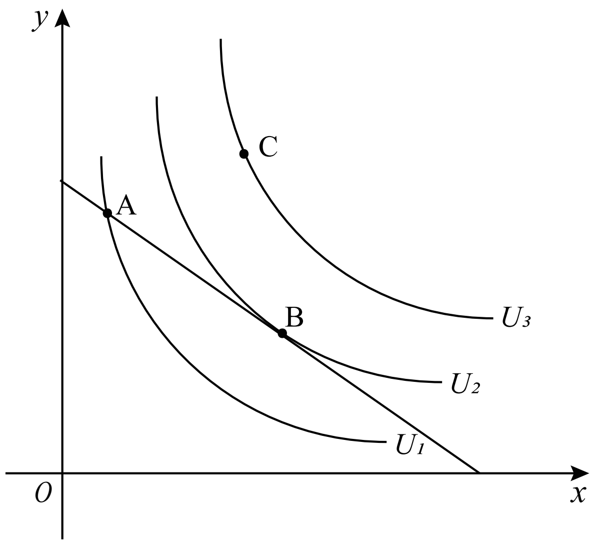

Figure 3.19 Consumer's optimal choice as an interior solution

Thus, at the interior optimal bundle, the following two condition must be satisfied:

(1) ***The slope of the indifference curve equals the slope of budget line***. Since the slope of indifference curve is $-MRS_{xy}$ and the slope of the budge line is $-\frac{p_x}{p_y}$, at B, then we have

$$MRS_{xy} = \frac{p_x}{p_y}. \tag{3.3}$$

(2) ***The choice is on the budget line***:

$$p_x x + p_y y = I. \tag{3.4}$$

In other words, at the point of satisfaction maximization, the MRS between the two goods equals the price ratio, and also satisfy the budget line. Thus, *equations (3.3) and (3.4) fully characterize the consumer' interior optimal choice*.

Note that, since $MRS_{xy} = \frac{MU_x}{MU_y}$, inserting it into the marginal equality condition (3.3), it can be rewritten as

$$\frac{MU_x}{MU_y} = \frac{p_x}{p_y}, \tag{3.5}$$

or equivalently,

$$\frac{MU_x}{p_x} = \frac{MU_y}{p_y}, \tag{3.6}$$

which means the marginal utility per dollar for x equals the marginal utility per dollar for y.

Example 3.8 (Example 3.5 continued again) Suppose that a consumer's utility function is given by

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

and the prices and income are given by

$$p_x = 2,\ p_y = 1,\ I = 100.$$

Substituting $MRS = \frac{y}{x}$, $p_x = 2$ and $p_y = 1$ into equations (3.3) and (3.4), we obtain

(1) $\frac{y}{x} = \frac{2}{1} \Rightarrow y = 2x$.

(2) $2x + y = 100$.

Substituting $y = 2x$ from (1) into the budget line (2), we have

$$2x + 2x = 100$$

and then

$$4x = 100.$$

Thus, we have $x^* = 25$ and $y^* = 50$.

Case 2: Corner solution (such as linear indifference curve): In this case, the marginal rate of substitution for any good is not equal to the slope of the budget line, and then the consumer maximizes satisfaction by consuming only one of the two goods as shown in the following example.

Example 3.9 (Example 3.4 continued again) Suppose $U = x + 2y$. Then $MRS = \frac{1}{2}$. The optimal consumption depends on the slopes of the budget line and the indifference curve.

i) $MRS_{xy} < \frac{p_x}{p_y} \Rightarrow$ Not consuming good x as depicted in Figure 3.20.

Let $p_x = 1$, $p_y = 1$, and $I = 20$. The budget line is $x + y = 20$, and its slope $= -1 \neq -MRS$. So y-intercept $= \frac{I}{p_y} = 20$. Thus $x^* = 0$ and $y^* = 20$.

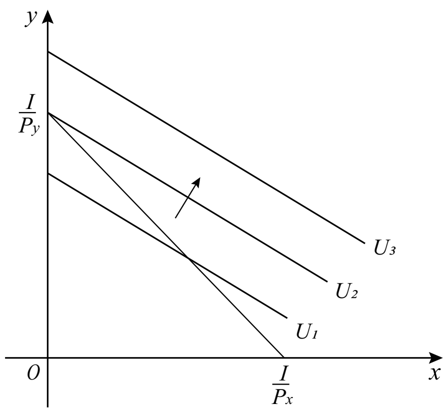

Figure 3.20 Corner solution: only good y matters

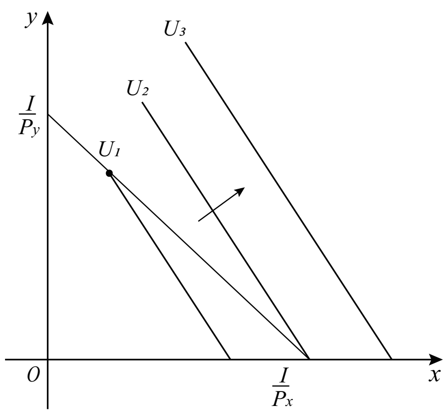

Figure 3.21 Corner solution: only good x matters

ii) $MRS_{xy} > \frac{p_x}{p_y} \Rightarrow$ Not consuming good y as depicted in Figure 3.21. Let $p_x = 1$ and $p_y = 3$. So $x^* = \frac{I}{p_x} = 20$ and $y^* = 0$.

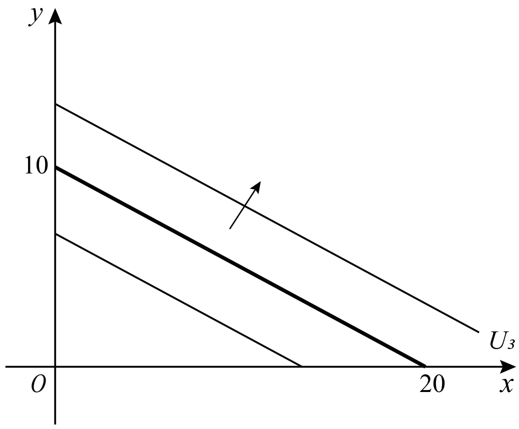

Figure 3.22 Indifference curves and budget line are parallel: infinitely many optimal choices

iii) $MRS_{xy} = \frac{p_x}{p_y} \Rightarrow$ All consumption bundles on the budget are optimal consumptions as depicted in Figure 3.22.

Let $p_x = 1$ and $p_y = 2$. The slopes of the budget line and the indifference curves are the same, any bundle (x^*, y^*) satisfying $x + 2y = 20$ is optimal.

3.4 The Composite-Good Convention

So far, our analysis is only for a two-good world, but the general principles can be applied to a world of many goods. Though many goods cannot be directly shown on a two-dimensional graph, it is possible to deal with multiple goods in two dimensions by treating a number of goods as a composite good.

Suppose there are many goods, $x, y, ..., z$. We can measure the consumption for x by treating other goods, $(y, ..., z)$, as a composite good. Consumption of the composite good is gauged by total outlays on it, in other words, total outlays on all goods other than x. Thus, all analysis and conclusions for a two-good world also hold for a world of many goods.

Notice that price of the composite is equal to one. Then slope of budget line $= -\frac{p_x}{1}$, and so $MRS = \frac{p_x}{1} = p_x$.

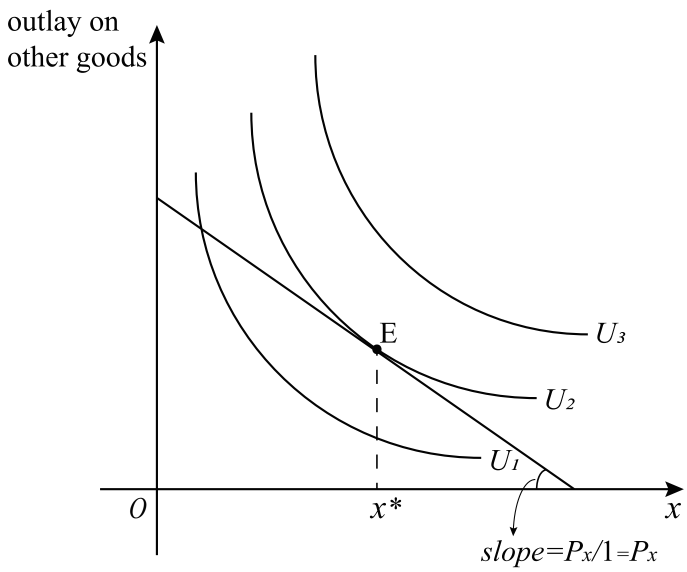

Figure 3.23 The composite-good convention: optimal choice with more than two goods

3.5 Revealed Preference

If a consumer chooses one market basket over another and the chosen market basket is more expensive than the alternative, then the consumer must prefer the chosen market basket.

If an individual facing budget line l_1 chose market basket A rather than market basket B, A is revealed to be preferred to B as depicted in Figure 3.24. Likewise, the individual facing budget line l_2 chooses market basket B, which is then revealed to be preferred to market basket D. Whereas A is preferred to all market baskets constrained by l_1 and l_2, all baskets greater than A are preferred to A.

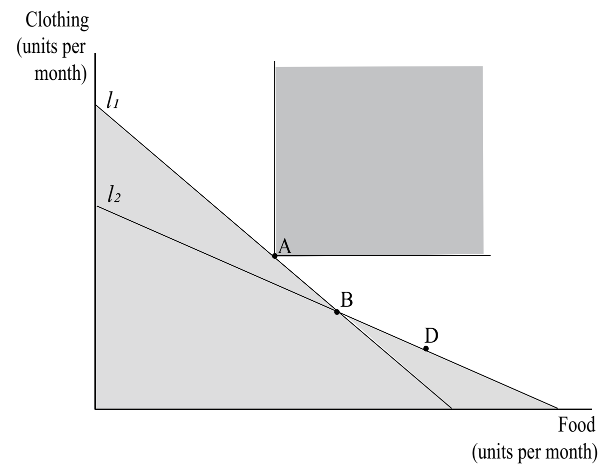

Figure 3.24 Revealed preference: Two budget lines

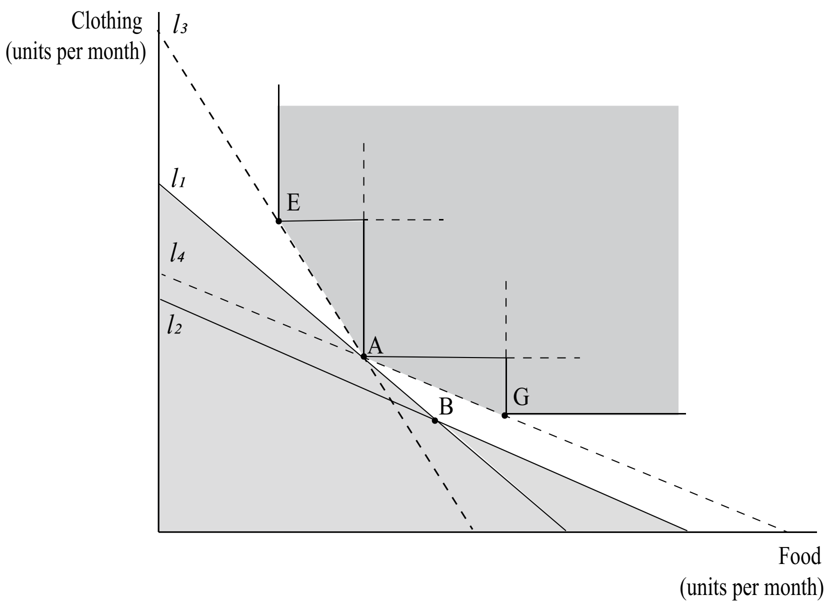

Figure 3.25 Revealed preference: Four budget lines

Facing budget line l_3, the consumer chooses E, which is revealed to be preferred to A (because A could have been chosen) as shown in Figure 3.25. Likewise, facing line l_4, the individual chooses G, which is also revealed to be preferred to A. Whereas A is preferred to all market baskets by l_1 and

l_2, all market baskets in the upper shaded area are preferred to A.

With more and more budget lines, consumer's preference is revealed.

Chapter 4

Individual and Market Demand

This chapter discusses the influences of income and price changes, and how to determine individual and market demand.

4.1 Effect of Income Changes

An increase in income, with the prices of all goods fixed, causes consumers to alter their choice of market baskets.

Income-consumption curve (ICC): Curve tracing the utility-maximizing combinations of two goods as a consumer's income changes.

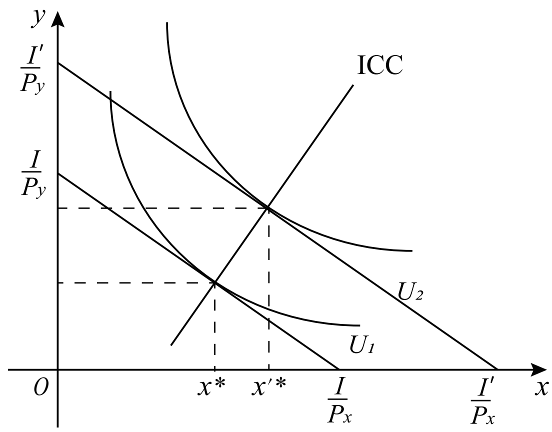

Figure 4.1 Tracing an income consumption curve

Thus, at each point along a price consumption curve, utility is maximized when all income is spent.

To find an income-consumption curve, we change I, leaving prices fixed. As shown in Figure 4.1, the demand of both x and y rises when income rises.

Normal good: An increase in a person's income can lead to more consumption of the good being purchased when p_x and p_y are fixed.

Most goods are normal goods, especially when income is low.

Inferior good: An increase in income can lead to less consumption of one of the two goods being purchased when p_x and p_y are fixed.

As income continuously increases, more and more goods become to be inferior goods such as potato, low quality goods, etc.

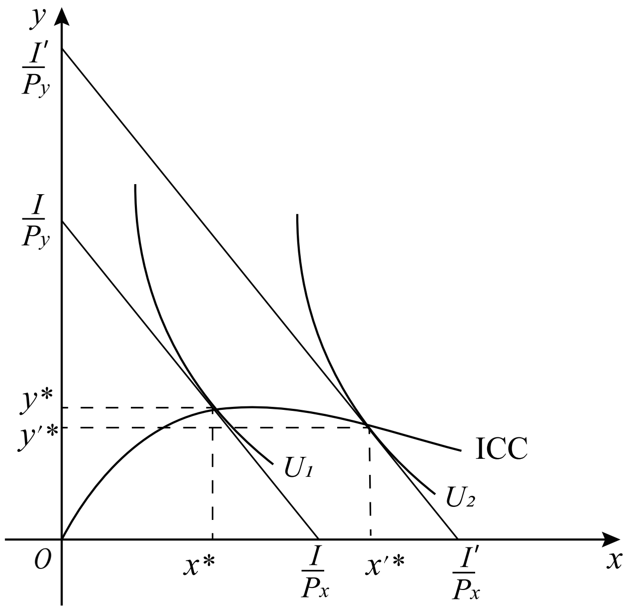

Figure 4.2 Income-consumption curve when good y is an inferior good

As shown in Figure 4.2, we have $I' > I$, $x'^* > x^*$, and $y'^* < y^*$. So, x is normal good and y is inferior good. This curve illustrates that, for all income levels, only x is normal. Notice that if there are only two goods in the economy, they cannot be inferior goods at the same time.

Engel curve: Curve relating the quantity of a good consumed to income.

In the left diagram in Figure 4.3, food is a normal good and the Engel curve is upward sloping.

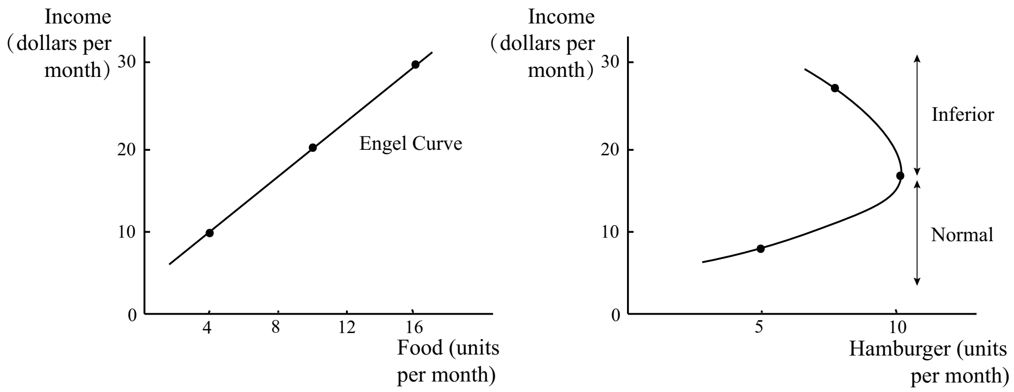

Figure 4.3 Engel cures

In the right diagram in Figure 4.3, however, hamburger is a normal good for income less than $18 per month and becomes an inferior good for income greater than $18 per month. That is, the income-consumption curve has a positive slope for low incomes, and then it takes a negative slope for even higher incomes.

Expenditures	Less than 19,999	10,000-29,999	20,000-39,999	30,000-49,999	40,000-59,999	50,000-69,999	70,000 and above
Entertainment	1,038	1,165	1,407	1,969	2,131	2,548	4,655
Owned Dwelling	1,770	2,134	2,795	3,581	4,198	5,566	11,606
Rented Dwelling	3,919	3,657	4,054	3,878	4,273	3,812	3,072
Health Care	1,434	2,319	3,124	3,539	3,709	4,702	6,417
Food	3,466	3,706	4,432	5,194	5,936	6,486	10,116
Clothing	798	766	960	1,321	1,518	1,602	2,928

Table 4.1: Annual U.S. Household Consumer Expenditures

Example 4.1 (Consumer Expenditures in the United States) We can derive Engel curves for groups of consumers. This information is particularly useful if one wants to see how consumer spending varies among different income groups.

From Table 4.1, one can see that health care and entertainment are normal goods, as expenditures increase with income. Rental housing, however, is an inferior good for incomes above $40,000.

The Supplemental Nutrition Assistance Program (SNAP)

The Supplemental Nutrition Assistance Program (SNAP), formerly known as the Food Stamp Program, is a federal program that provides food-purchasing assistance for low- and no-income people. Under SNAP roughly 1 in 7 Americans receive aid across. If people only have SNAP funds, they can't withdraw cash or use their cards to pay for housing or non-food items. What effects are this program?

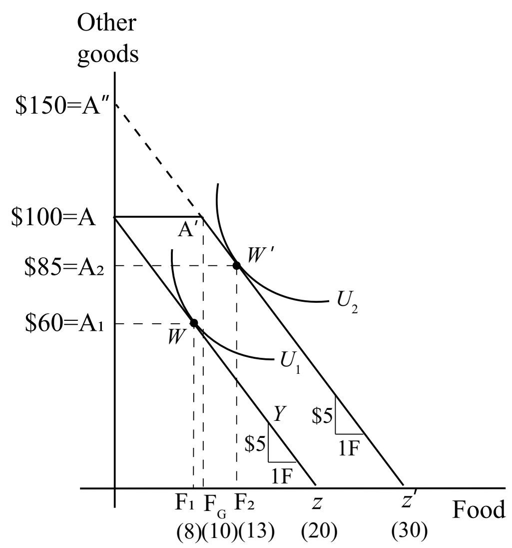

Figure 4.4 Evaluating SNAP using consumer theory

Consumer theory can be used to evaluate this program. We show this by considering a specific example in which a consumer receives \$50 worth of SNAP funds a week. We assume the consumer has a weekly income of \$100 and the price of food is $p_f = \$5$ per unit.

The pre-subsidy budget line is AZ. The SNAP subsidy shifts the budget line to $AA'Z'$. Over the AA' range, the budget line is horizontal since the \$50 in free good SNAP funds permits the recipient to purchase up to 10 units of good while leaving his or her entire income of \$100 to be spent on other goods. Over 10 units of food, the consumer has to pay for it by \$5 per unit. Thus, the $A'Z'$ portion of the budget line has a slope of -5. Note that this new budget line is not straight line, but has a kink at A'.

The SNAP funds will affect the recipient in one of two ways. Figure 4.4 shows a possibility. If the consumer spends more than \$50 on food, the equilibrium, W', occurs on the $A'Z'$ portion of the budget line. The consequences of the SNAP funds are exactly the same as when the consumer receives a cash grant of \$50, leading to the budget line $A''Z$.

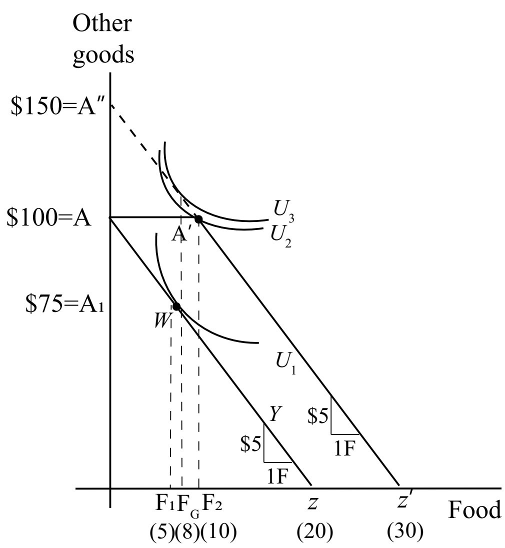

Figure 4.5 Another illustration of the effect of SNAP

Figure 4.5 shows another possible outcome of the SNAP subsidy. With a direct cash grant of \$50, the consumer would be better off on the indifference curve U_3, which is prohibited by the SNAP funds. The consumer has to choose among the options on the $AA'Z'$ budget line, and thus the best choice is the kink point A'. Therefore, the consumer would be better off if the subsidy is given as cash instead of as the SNAP funds.

In summary, in the first case, the consumer is equally well off under either cash grants or SNAP funds; in the second case, the consumer would be better off under cash grants. There is no such a case, however, where the consumer is better off with SNAP funds.

The same arguments and figures could be used to discuss a college trust fund. When given a college trust fund that must be spent on education, the student moves from W to A', a corner solution. If, however, the trust fund could be spent on other consumption as well as education, the student would be better off on the indifference curve U_3.

4.2 Effect of Price Changes: Deriving the Demand Curve

An increase in price of a good, with the prices of all other goods and income fixed, can also causes consumers to alter their choice of market baskets.

Price-consumption curve: Curve tracing the utility-maximizing combinations of two goods as the price of one changes.

Individual demand curve: Curve relating the quantity of a good that a single consumer will buy to its price.

There are two approaches to derive the **demand curve** of a consumption good from utility maximization.

1. Graphical derivation

To find the demand for x, we change p_x, while holding p_y and I constant.

Let $p_x > p_x' > p_x''$. We can get the baskets that maximize utility for these prices and have other optimal baskets in a similar way. Connecting these points smoothly, we can get the **price-consumption curve** along which *utility is maximized and all income is spent*, as depicted in the upper diagram of the Figure 4.6.

The lower diagram in Figure 4.6 gives the demand curve, which relates the price of the good to the quantity demanded Thus, **demand** for x is a function of p_x, p_y and I.

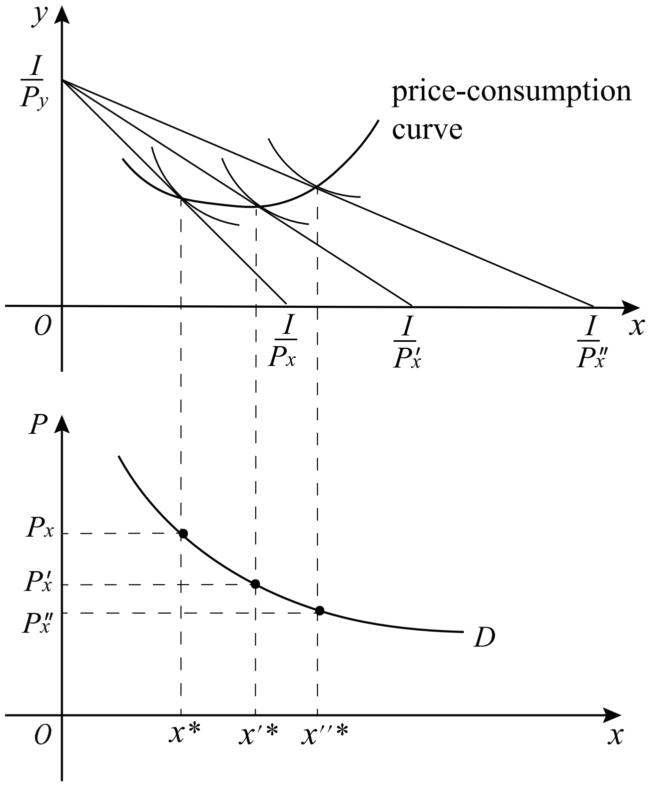

Figure 4.6 Deriving demand curve from utility maximization

2. Algebraic derivation

Use the two conditions for utility maximization:

$$MRS = \frac{p_x}{p_y}, \tag{4.1}$$

and

$$p_x x + p_y y = I. \tag{4.2}$$

Solving equation (4.1) and (4.2), we can find x^* and y^* as functions of p_x, p_y and I. Let illustrate with the following example.

Example 4.2 Suppose that preferences are represented by Cobb-Douglas utility function

$$U(x, y) = x^{\frac{1}{3}} y^{\frac{2}{3}}.$$

Then,

$$MRS = \frac{y}{2x}.$$

Equalizing MRS and $\frac{p_x}{p_y}$, we have

$$\frac{y}{2x} = \frac{p_x}{p_y},$$

and then

$$y = \frac{2xp_x}{p_y}$$

.

Substituting $y = \frac{2xp_x}{p_y}$ into the budget line, we have

$$p_x x + p_y(\frac{2xp_x}{p_y}) = I$$

or

$$3p_x x = I.$$

Therefore, we obtain

$$x^* = \frac{I}{3p_x} = D_x.$$

Let $I = 50$ and $p_y = 2$. Then $x^* = \frac{50}{3p_x}$.

To derive the demand curve, consider the following numerical illustration:

- $p_x = 1 \Rightarrow x = \frac{50}{3} = 16.6,$

- $p_x = 5 \Rightarrow x = \frac{50}{15} = \frac{10}{3} = 3.3$,
- $p_x = 10 \Rightarrow x = \frac{50}{30} = \frac{5}{3} = 1.6$.

Connecting these points and other points obtained in a similar way s-moothly, we can have an indifference depicted in Figure 4.7.

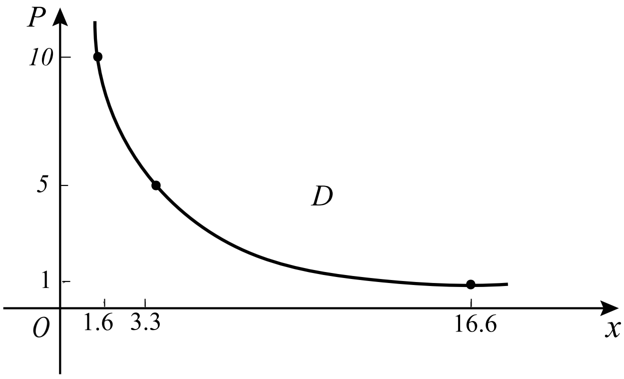

Figure 4.7 A downward-sloping and convex-to-origin demand curve

For a general Cobb-Douglas utility function:

$$u(x,y) = x^a y^{1-a}, \quad 0 < a < 1,$$

by the same way, we can derive the demand functions for x and y:

$$x(p_x, p_y, I) = \frac{aI}{p_x}$$

and

$$y(p_x, p_y, I) = \frac{(1-a)I}{p_y}.$$

4.3 Income and Substitution Effects of a Price Change

Any two goods could exhibit the following relationships:

- Two goods are **substitutes** if an increase in the price of one leads to an increase in the demand of the other.

- Two goods are **complements** if an increase in the price of one good leads to a decrease in the demand of the other.

- Two goods are **independent** if a change in the price of one good has no effect on the demand of the other.

The fact that goods can be complements or substitutes suggests that when studying the effects of price changes in one market, it may be important to look at the consequences in related markets.

- **Substitution effect**: Change in consumption of a good associated with a change in its price, with ***the level of utility held constant***.

  Say, when the price of a good falls, consumers will tend to buy more of the good that has become cheaper and less of those goods that are now relatively more expensive.

- **Income effect**: Change in consumption of a good resulting from a change in purchasing power, with ***prices held constant***.

  Say, when the price of a good falls, because it is now cheaper, consumers enjoy an increase in real purchasing power. The change in demand resulting from this change in real purchasing power is the income effect.

The total effect of a change in price is given theoretically by the sum of the substitution effect and the income effect:

$$\text{Total effect} = \text{substitution effect} + \text{income effect}$$

Income and Substitute Effects: Normal Good

In Figure 4.8, the consumer is initially at A on budget line RS and U_1. With a decrease in the price of food, the consumer moves to B on the new budget line RT and U_2. When the price of food falls, consumption increases by F_1F_2 as the consumer moves to B. The resulting change in food purchased can be broken down into substitution and income effects.

Since the real purchasing power increases due to the price decrease, to find the substitution effect, the consumer should spend less income in order to keep the same utility level as U_1. Note that any parallel line of RT has the same slope.

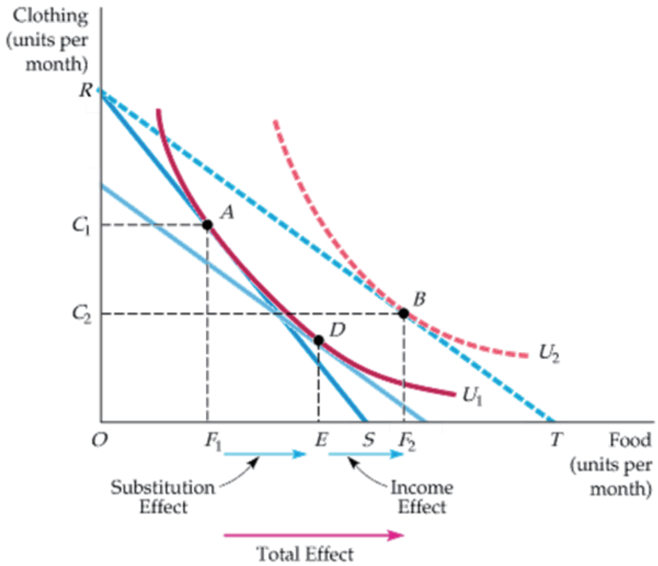

Figure 4.8 Income and substitution effects: normal good

Thus, the way to determine the substitution effect (and consequently income effect) is simply as follows:

Draw a parallel line of the new budget line RT***, which is tangent to the initial indifference curve*** U_1.

That is, we shift the new budget line RT parallelly toward the initial indifference curve U_1 and stop at the tangent point D on U_1. Then, moving from A to D on U_1 represents the effect of the decrease in the food price on food consumption, which is F_1E, while maintaining the same level of satisfaction. This is the **substitution effect**.

The remaining change, from D on U_1 to B on U_2 , is the **income effect**. The difference between the budget line RT and its parallel budget line is EF_2 that represents the effect of the increase in real purchasing power on food consumption, resulting in the increase in utility from U_1 to U_2.

Because food is a normal good, the income effect EF_2 is positive, hence their income and substitute effects move in the same direction.

Income and Substitute Effects: Inferior Good

In Figure 4.9, the consumer is initially at A on budget line RS. With a decrease in the price of food, the consumer moves to B. Again, the resulting change in food purchased can be broken down into a substitution effect, F_1E (associated with a move from A to D), and an income effect, EF_2 (associated with a move from D to B). In this case, food is an inferior good since the income effect is negative. Thus, the ***income and substitute effects move in the opposite directions for inferior good***.

However, *since the substitution effect dominates the income effect, the decrease in the price of food leads to an increase in the quantity of food demanded.*

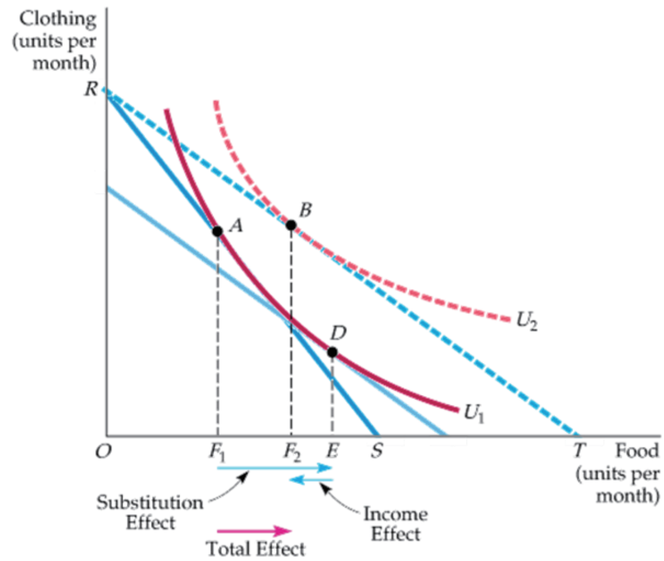

Figure 4.9 Income and substitution effects: inferior good

Income and Substitute Effects: Giffen Good

So far, the demand curves we derived are all downward-sloping, that is, demand for x will decrease as its price increases. However, it is possible for a consumer to have indifference curves so that the law of demand does not hold for some good.

A good is said to be a **Giffen good** if the quantity demanded will fall when its price falls.

What kind of goods can be Giffen goods? The Giffen good is rarely of practical interest because it requires a large negative income effect. But, in practice, the income effect is usually small.

As shown in Figure 4.10, a lower price could lead to less consumption. The consumer purchases less of food when its price falls. Note that the indifference curves that produce this result are downward-sloping, non-intersecting, and convex; that is, they do not contradict any of the basic assumptions about preferences.

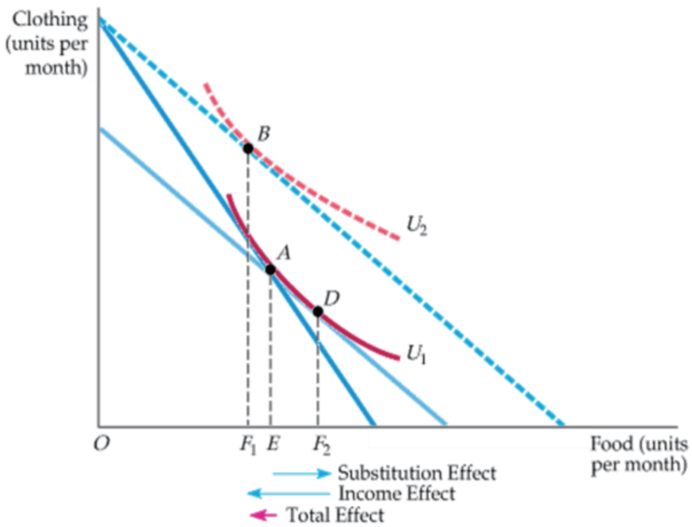

Figure 4.10 Upward-sloping demand curve: the Giffen good

Because the income effect F_2F_1 is negative (so that substitution effect and income effect move in the opposite directions,) and dominates the substitution effect EF_2, the decrease in the price of food leads to a lower quantity of food demanded.

Thus, a Giffen good is fully characterized by the following two features:

(1) it must be inferior good;

(2) its income effect dominates its substitution effect (in absolute value).

That is, ***a Giffen good is the special subset of inferior goods in which the income effect dominates the substitution effect***.

Therefore, *a Giffen good must be an inferior good, but an inferior good may not be a Giffen good.*

4.4 From Individual to Market Demand

We have derived the demand curve for an individual consumer. To obtain a market demand curve, we simply add all individual demands at given prices. Two points should be noted:

(1) The market demand curve will shift to the right as more consumers enter the market.

(2) Factors that influence the demands of many consumers will also affect market demand.

The ***aggregation of individual demands*** into market becomes important in practice when market demands are built up from the demands of different demographic groups or from consumers located in different areas.

Example 4.3 Consider the numerical example with two consumers. Making the summation of two individuals' demand, we obtain the market demand as shown in Table 4.3 and Figure 4.11.

p_x for good x	D_x^A for A	D_x^B for B	Market demand D_x
\$1	10	12	22
\$2	7	5	12
\$3	3	1	4

Table 4.2: Determining the market of demand curve

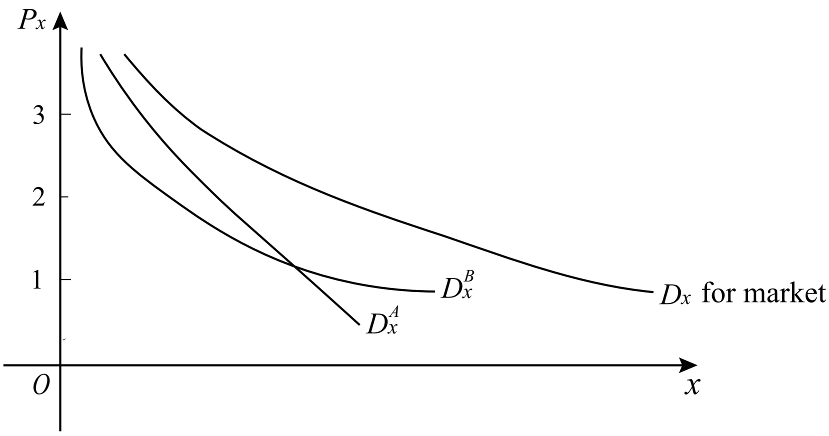

Figure 4.11 From individual demand to market demand

4.5 Consumer Surplus

Consumers purchase goods and services because they are better off from doing so; otherwise, the purchase would not take place. The term **consumer surplus** refers to the net benefit, or gain.

Consumer surplus: Difference between what a consumer is willing to pay for a good and the amount actually paid.

To obtain the measure of consumer surplus associated with purchasing a certain amount of a good or service, we may ask the question: What is the maximum amount you would be willing to pay, that is, what is the maximum total benefit, and what is the total cost? After we have the answers, the consumer surplus is defined by:

Consumer surplus = Total benefit - Total cost.

In other words, consumer surplus is the difference between the willingness-to-pay of the consumer and what the person actually has to pay for it.

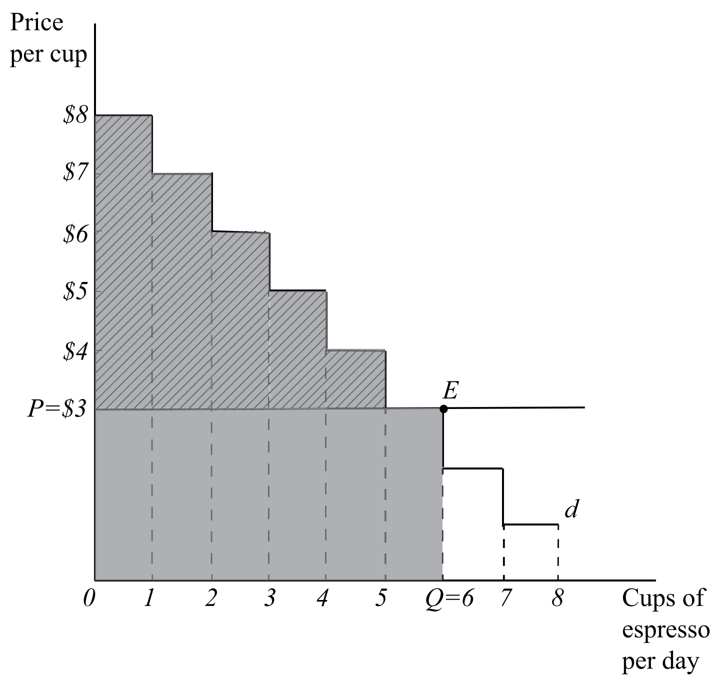

Figure 4.12 The discrete approach of calculating consumer surplus

To show this, let us consider a specific example. Suppose that a person purchases 6 cups of espresso at the price $p = \$3$. The total cost $= 3 \times 6 = 18$. Total benefit from purchasing 6 units at a price \$ 3, however, is the sum of marginal benefit (the sum of 6 units shaded rectangles of Figure 4.12), i.e.,

$$\text{total benefit} = \$8 + \$7 + \$6 + \$5 + \$4 + \$3 = 33.$$

Thus, the consumer surplus = total benefit − total cost $= 33 - 18 = 15$.

In general, with a smooth demand curve indicated in Figure 4.13, consumer surplus equals the area TEP.

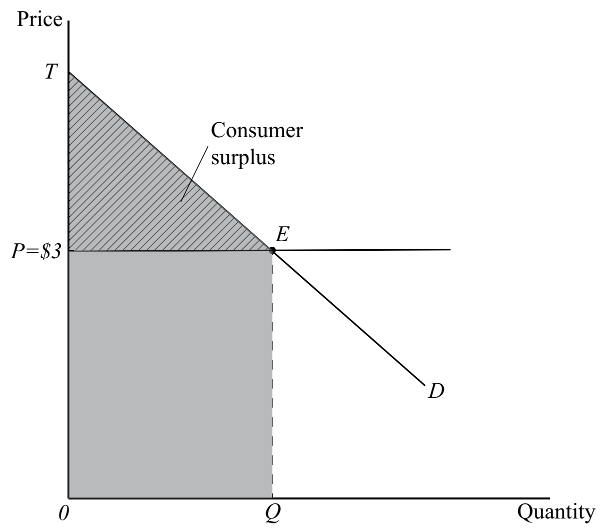

Figure 4.13 Consumer surplus under a downward-sloping continuous demand curve

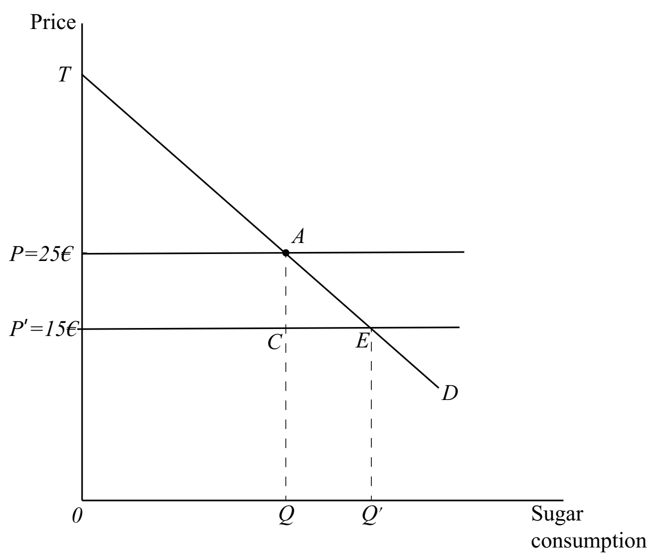

Figure 4.14 The effect of a price reduction on consumer surplus

The concept of consumer surplus can also be used to identify the net benefit of a change in the price of a commodity. In Figure 4.14, at a price of 25 cents per unit, consumer surplus is TAP. At a price of 15 cents per unit, consumer surplus is TEP'. The increase in consumer surplus from the price reduction is thus the shaded area $PAEP'$, which is a measure of the benefit to consumers of a reduction in the price from 25 to 15 cents.

Consumer surplus measures the aggregate benefit that consumers obtain from buying goods in a market. When we combine consumer surplus with the aggregate producer surplus that producers obtain, we can evaluate both the costs and benefits not only of alternative market structures, but of public policies that alter the behavior of consumers and firms in those markets.

4.6 Network Externalities

Network externality: When each individual's demand depends on the purchases of other individuals.

Positive network externality: The quantity of a good demanded by a typical consumer increases in response to the growth in purchases of other consumers (e.g., social network—telephones, emails, Facebook, Tik Tok, WeChat; toys and fads etc.).

Negative network externality: The quantity demanded decreases to the growth in purchases of other consumers (e.g., an individual has the desire to own exclusive or unique goods).

Bandwagon Effect

The existence of positive network externalities gives rise to Bandwagon effect.

Bandwagon effect: Positive network externality in which a consumer wishes to possess a good in part because others do, i.e., to have a good because of the perception that almost everyone else has it.

With a positive network externality, the quantity of a good that an individual demands grows in response to the growth of purchases by other individuals. Here, as the price of the product falls from \$30 to \$20, the bandwagon effect causes the demand for the good to shift to the right, from D_{40} to D_{80}.

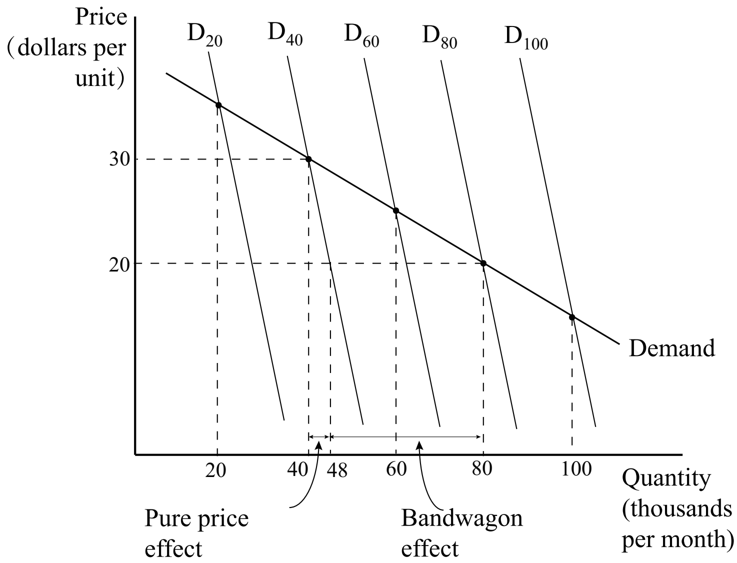

Figure 4.15 Positive network externalities: Bandwagon effect

Year	Facebook (Meta) Users (Millions)	Hours Per User Per Month
2004	1	—-
2005	5.5	—-
2006	12	<1
2007	50	2
2008	100	3
2009	350	5.5
2010	500	7
2011	766	7.5
2012	980	8.5
2013	1171	9
2014	1334	10
2015	1517	10.5
2016	1654	11

Table 4.3: Facebook Users

Positive network externalities have been crucial drivers for many modern technologies over many years.

Example 4.4 (Facebook (Meta)) By early 2011, with over 700 million users, Meta previously known as Facebook became the world's second most visited website (after Google). By 2022, there are about 2 billion daily users. A strong positive network externality was central to Facebook's success.

Snob Effect

The existence of negative network externalities gives rise to snob effect. **Snob effect**: Negative network externality in which a consumer wishes to own an exclusive or unique good.

The quantity demanded of a "snob" good is higher when the fewer is the number of people who own it (e.g., garments specially designed by a well-known fashion designer, special designs of watches, automobiles, buildings, etc.).

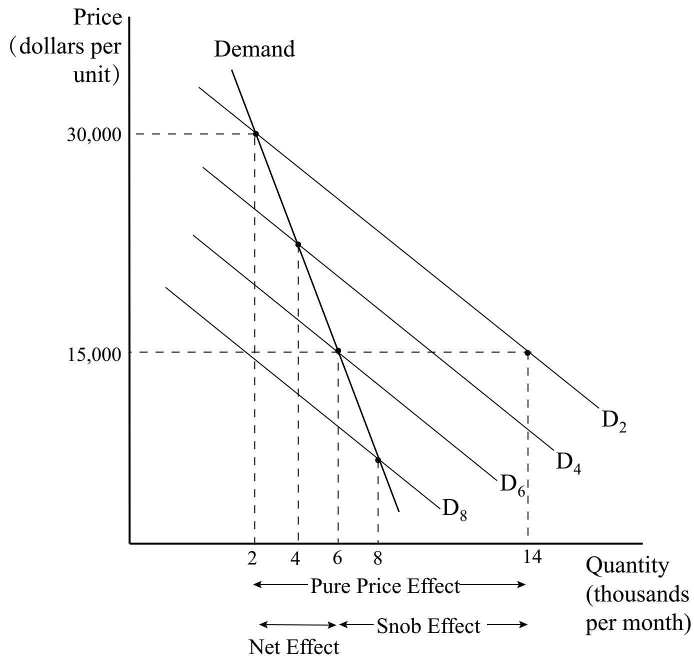

Figure 4.16 Negative network externalities: Snob effect

Thus, snob effect is a negative network externality in which the quantity of a good that an individual demands falls in response to the growth of purchases by other individuals.

In Figure 4.16, as the price falls from \$30,000 to \$15,000 and more people buy the good, the snob effect causes the demand for the good to shift to the left, from D_2 to D_6.

Part III

Supply Side of Market

Part 2 presents the theoretical core of production and competitive firm's supply and market supply.

Chapter 6

Theory of Production

In this chapter and the next we discuss the theory of the firm, which describes how inputs (such as labor, capital, and raw materials) can be transformed into outputs, how a firm makes cost-minimizing production decisions, and how the firm's resulting cost varies with its output.

6.1 The Technology of Production

Firm: any organization that engages in production.

The production decisions of firms can be understood in three steps:

1. **Production Technology**: How inputs (also known as the factors of production) can be transformed into outputs.
2. **Cost Constraints**: What possible combinations of inputs can be used with a given expenditure on production.
3. **Input Choices**: Just as a consumer is constrained by a limited budget, the firm is concerned about the cost of production and the prices of labor, capital, and other inputs.

We can divide inputs into the broad categories of labor, materials and capital, each of which might include more narrow subdivisions.

- ***Labor inputs*** include skilled workers (financial analyst, carpenters, engineers) and unskilled workers (cleaners, manual workers), as well as the entrepreneurial efforts of the firm's managers.

- ***Raw materials*** include steel, plastics, electricity, water, and any other goods that the firm buys and transforms into final products.

- ***Capital*** includes land, buildings, machinery and other equipment, as well as inventories.

Production Function: A relationship between inputs and outputs that identifies the ***maximum output*** produced by a firm for every combination of inputs with a given technology.

A production is **technologically efficient** if the maximum quantity of a commodity can be produced by *each* specific combination of inputs. Thus, *production specified by production function is technologically efficient*. If a firm is rational, it should always operate in a technologically efficient way to produce outputs.

We can present a production function in tabular, graphical, or mathematical form. For example, in mathematical form, we have

$$q = F(K, L),$$

where q is the number of units of output, K the number of units of capital input, L the number of units of labor input, and F is production function.

A common production function used as an example in economics is the so-called **Cobb-Douglas production function**:

$$F(K, L) = AK^{\alpha}L^{\beta},$$

where A is a parameter that can be interpreted as total productivity.

Suppose that $\alpha = \beta = \frac{1}{2}$ and $A = 1$. We have

$$q = K^{\frac{1}{2}}L^{\frac{1}{2}}.$$

Then, for a particular combination of inputs, say, $K = 16$ and $L = 36$, the maximum output obtainable with this technology is

$$q = (16)^{\frac{1}{2}}(36)^{\frac{1}{2}} = 4 \times 6 = 24.$$

6.2 Production in Short-Run

We first define the relevant concepts as follows.

Short run: A period of time in which changing the employment levels of some inputs is impractical, i.e., it is a time period in which at least one input is fixed.

Fixed inputs: Production factors that cannot be varied over the time period involved.

Long run: A period of time in which the firm can vary all its inputs.

For example, for a factory, short run is long enough to hire an extra worker but is not long enough to build an extra production line.

6.2.1 Various Product Curves in Short Run

Assume there are two factors of production: labor (L) and capital (K).

In the short run we usually assume K is fixed and L is a variable input. Given fixed K, the firm can vary L to produce different amounts of output.

- **Total Product (TP)**: total output produced;
- **Average Product (AP)**: Output per unit of a variable factor. That is, $AP =$ total product/variable factor.

  For example, average product of labor is given by

  $$AP_L = \frac{TP}{L}.$$

- **Marginal Product (MP)**: the change in output resulting from a change in the amount of the input, holding the quantities of other inputs constant. For example, marginal product of labor:

$$MP_L = \Delta TP/\Delta L.$$

The Average-Marginal Products Relationship

- If $MP_L > AP_L$, then AP_L must rise.
- If $MP_L < AP_L$, then AP_L must fall.
- If $MP_L = AP_L$, then AP_L reaches its maximum.

Note that if MP_L is constant, AP_L is identical to MP_L, hence constant in L.

Example 6.1 Suppose $q = F(K, L) = 10KL^2$. Then $AP_L = \frac{q}{L} = 10KL$ and $MP_L = \partial q/\partial L = 20KL$. It can be verified that $MP_L > AP_L$ implies AP_L increases.

Example 6.2 (Production with One Variable Input (Labor)) Consider the numerical example depicted in Table 6.1.

Fixed amount of land	Amount of labor (L)	Total product (TP_L)	Average product of labor (AP_L)	Marginal product of labor (MP_L)
5	0	0	$--$	$--$
5	1	50	50	50
5	2	150	75	100
5	3	300	100	150
5	4	400	100	100
5	5	480	96	80
5	6	540	90	60
5	7	580	83	40
5	8	610	76	30
5	9	610	68	0
5	10	580	58	-30

Table 6.1: Production with one variable input (labor)

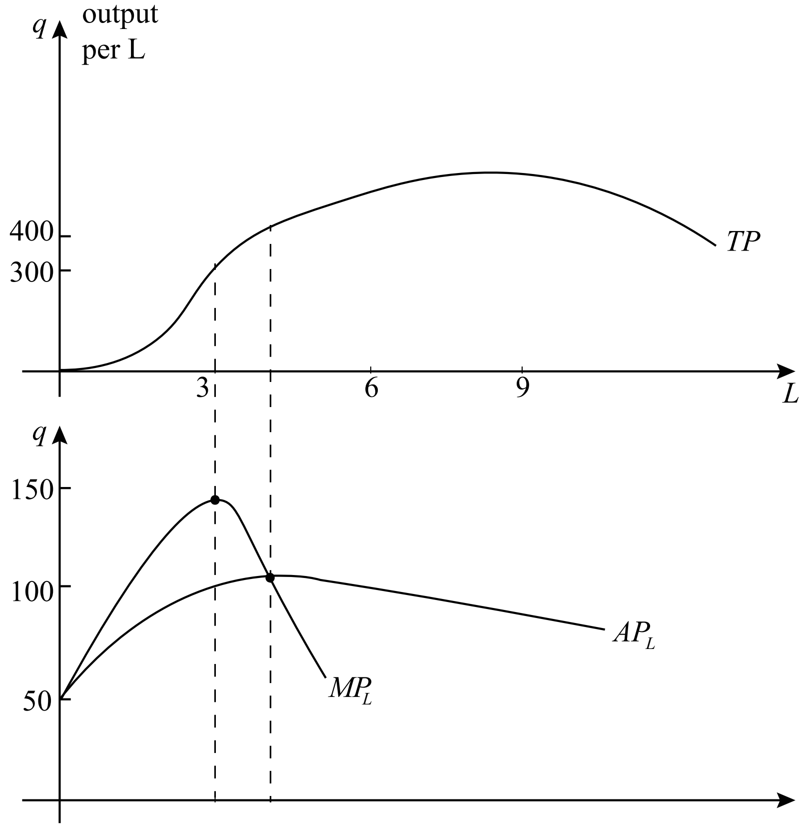

Figure 6.1 The relationship between TP, MP and AP

The total output curve in the upper diagram in Figure 6.1 shows the output produced for different amounts of labor input.

The average and marginal products in the lower diagram in Figure 6.1 can be obtained (using the data in Table 6.1) from the total product curve.

With 3 units of labor, the marginal product is 150. The average product of labor, however, is 100, and then the average product of labor continue to increase. With 4 units of labor, the marginal product of labor and the average product of labor are interest, the average product of labor reaches its maximum at this point.

When the labor input is 5 units, the marginal product is below the average product, so the average product is falling. Once the labor input exceeds 9 units, the marginal product becomes negative, so that total output falls as more labor is added. When total output is maximized, the slope of the tangent to the total product curve is 0, as is the marginal product. Beyond that point, the marginal product becomes negative.

6.2.2 The Geometry of Production Curves

(1) ***Average product (*$AP_L = \frac{q}{L}$*) is the slope of the line which connects the origin to the* TP *curve.***

This is because the slope of any line from the origin to a point on the TP curve has slope $= \frac{\Delta q}{\Delta L} = \frac{q-0}{L-0} = \frac{q}{L}$. At point C in Figure 6.3, AP_L reaches its maximum since the ray OC is the steepest ray from the origin that still touches the TP curve.

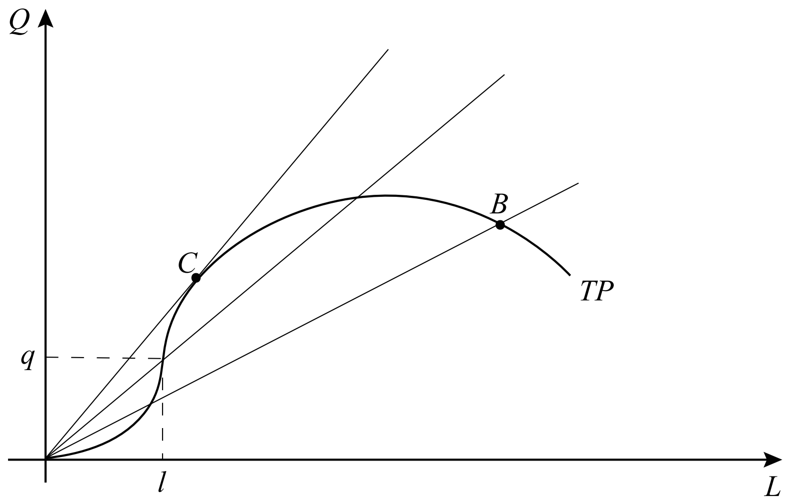

Figure 6.2 The geometry of production curves

(2) ***Marginal product (*$MP_L = \frac{\Delta q}{\Delta L}$*) is the slope of the* TP *curve.***

The following law of diminishing marginal product (DMP) reveals the shape of the TP curve. MP_L increases to a point and then decreases. This is, the slope of TP first increases and then eventually starts to decrease. When $MP_L = 0$, the TP curve reaches its maximum.

6.2.3 The Law of Diminishing Marginal Returns

The law of diminishing (marginal) returns (DMR): As the amount of a variable input increases with *a fixed amount of other inputs and fixed technology*, a point is reached beyond which the marginal product of the variable input begins to fall.

In short, ***the marginal product of a variable input eventually decreases with a given technology and other inputs fixed.***

Example 6.3 Farm adds fertilizer (variable) to an acre of land (fixed); and the numerical example as follows:

K,L	q	MP_L	Diminishing Returns
1,0	0	–	
1,1	10	10	
1,2	30	20	$\downarrow$
1,3	40	10	$\downarrow$
1,4	45	5	$\downarrow$
1,5	48	3	$\downarrow$

Table 6.2: Production with one variable input (labor)

Remarks on DMR

Technology Improvement. Labor productivity (output per unit of labor for an economy or industry as a whole) can increase if there are improvements in technology, even though any given production process exhibits diminishing returns to labor. The law of diminishing marginal returns was central to the thinking of political economist Thomas Malthus (1766-1834). Malthus predicted that as both the marginal and average productivity of labor fell and there were more mouths to feed, mass hunger and starvation would result. Malthus was wrong (although he was right about the diminishing marginal returns to labor). Over the past century, technological improvements have dramatically altered food production in most countries (including developing countries, such as India). As a result, the average product of labor and total food output have increased. Hunger remains a severe problem in some areas, in part because of the low productivity of labor there.

Example 6.4 (Malthus and The Food Crisis) Table 6.3 shows that technological improvements have dramatically altered food production.

Year	Index of World Food Production Per Cap
1961-64	100
1965	101
1970	105
1975	106
1980	109
1985	115
1990	117
1995	119
2000	127
2005	135
2010	146
2013	151

Table 6.3: Index of world food production per cap.

6.2.4 Labor Productivity and the Standard of Living

Consumers in the aggregate can increase their rate of consumption in the long run only by increasing the total amount they produce. Understanding the causes of productivity growth is an important area of research in economics.

Development of new technologies allowing factors of production to be used more effectively.

Example 6.5 (Labor Productivity and the Standard of Living) Will the standard of living in the United States, Europe, and Japan continue to improve, or will these economies barely keep future generations from being worse off than they are today?

Because the real incomes of consumers in these countries increase only as fast as productivity does, the answer depends on the labor productivity of workers.

	United States	Japan	France	Germany	UK
GDP Per Hour	\$62.41	\$39.39	\$60.28	\$58.92	\$47.39
Years	Labor (%) Productivity	Labor (%) Productivity	Labor (%) Productivity	Labor (%) Productivity	Labor (%) Productivity
1970-1979	1.7	4.5	4.3	4.1	3.2
1980-1989	1.4	3.8	2.9	2.1	2.2
1990-1999	1.7	2.4	2.0	2.3	2.3
2000-2009	2.1	1.3	1.2	1.1	1.5
2010-2014	0.7	1.2	0.9	1.2	0.5

Table 6.4: Annual rate of growth of labor productivity (in 2010 U.S dollars) in developed countries

6.3 Production in the Long-Run

Now we return to the case of long-run and assume that labor and capital both are variable, and introduce the important concept of isoquant in production decision.

6.3.1 Production Isoquants

Production Isoquant: A curve that shows all the combinations of inputs that, when used in a technologically efficient way, yield the same total output.

Isoquants show the **input flexibility** that firms have when making production decisions: They can usually obtain a particular output by substituting one input for another. It is important for managers to understand the nature of this flexibility.

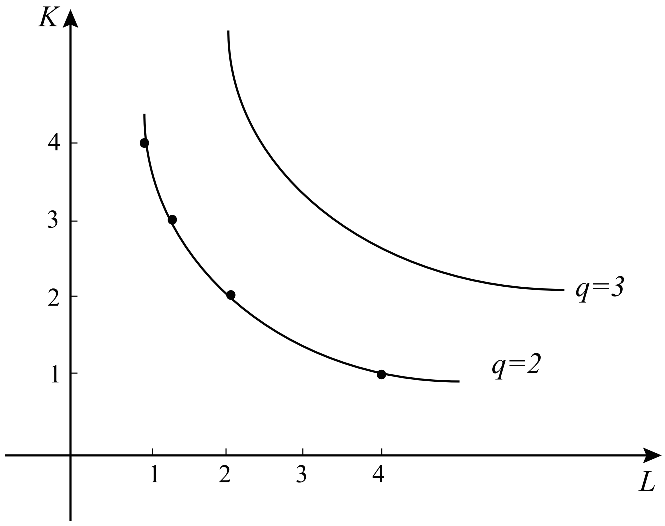

Figure 6.3 A set of production isoquants or an isoquant map

For example, if $q = K^{\frac{1}{2}}L^{\frac{1}{2}}$, the isoquant for $q = 2$ must include the following pair of inputs:

$$K = 1, L = 4;$$
$$K = 2, L = 2;$$
$$K = 3, L = \frac{4}{3};$$
$$K = 4, L = 1.$$

Isoquants are very similar to indifference curves in their characteristics. By analogous reasoning, we can explain several characteristics of isoquants.

i) ***Isoquant slopes downward.*** If we increase the quantity of one input employed and keep output unchanged, then we must reduce the amount of the other inputs.

ii) ***Isoquants can never intersect***.

iii) ***Lying further to the northeast identify higher levels of outputs***.

iv) ***Isoquants are generally convex to the origin***.

6.3.2 Marginal Rate of Technical Substitution

The slope of an isoquant measures the marginal rate of technical Substitution between the inputs.

Marginal rate of technical substitution (MRTS): The rate at which one input can be substituted for another with output remaining constant.

By the same reasoning as indifference curves, the marginal rate of technical substitution is equal the absolute value of the slope of an isoquant:

$$MRTS_{LK} = \left|\frac{\Delta K}{\Delta L}\right|.$$

Also, it is equal to the ratio of the marginal products of the inputs, i.e.,

$$MRTS_{LK} = \frac{MP_L}{MP_K}.$$

An isoquants is strictly convex implies the marginal rate of technical substitution diminishes, known as the **diminishing marginal rate of technical substitution**.

Limiting Isoquants:

- **Perfect substitutes**: The rate at which capital and labor can be substituted for each other is the same no matter what level of inputs is being used.

  That is, isoquants are straight lines so that the MRTS is constant, and thus it is known as the **linear production function**, which can be written as

$$q = F(K, L) = aK + bL.$$

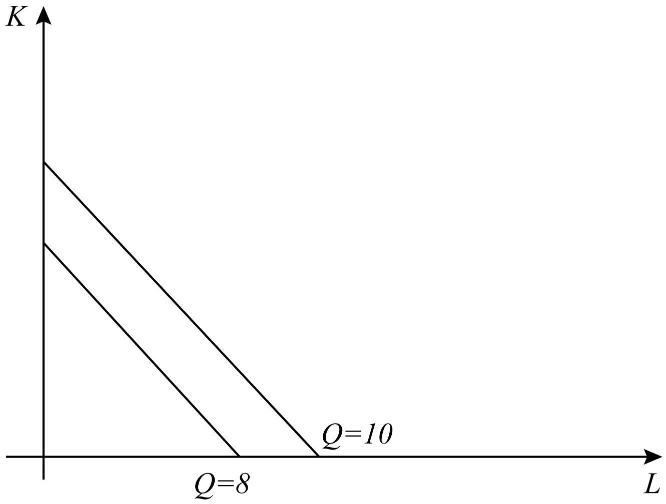

Figure 6.4 Perfect substitutes production function

- **Perfect complement**: Indicate that capital and labor cannot be substituted for each other in production.

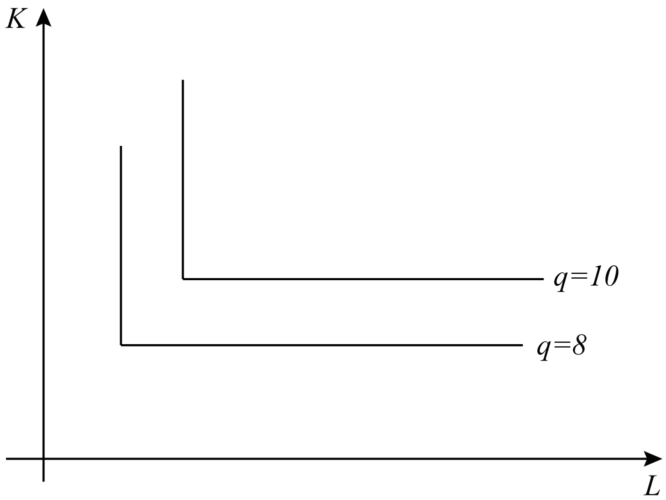

Figure 6.5 Production function with L-shaped isoquants

A production function with perfect complement is said to be the **fixed-proportions production function** (also known as the **Leonitief production function**), which can be written as

$$q = F(K, L) = \min\{aK, bL\}.$$

The isoquants for these two production functions are convex, but not strictly convex.

6.3.3 Returns to Scale

A firm can increase its output by increasing all inputs. We then want to know the scale of return.

Returns to scale: Effect on output of equal proportionate change in all inputs.

E.g., double inputs $\Rightarrow$ what will happen to output?

There are three types of production processes:

- **Constant returns to scale** (CRS): output doubles when all inputs are doubled;
- **Increasing returns to scale** (IRS): output more than doubles when all inputs are doubled;
- **Decreasing returns to scale** (DRS): output less than doubles when all inputs are doubled.

Graphically, the three types of returns to scale are illustrated in Figure 6.6.

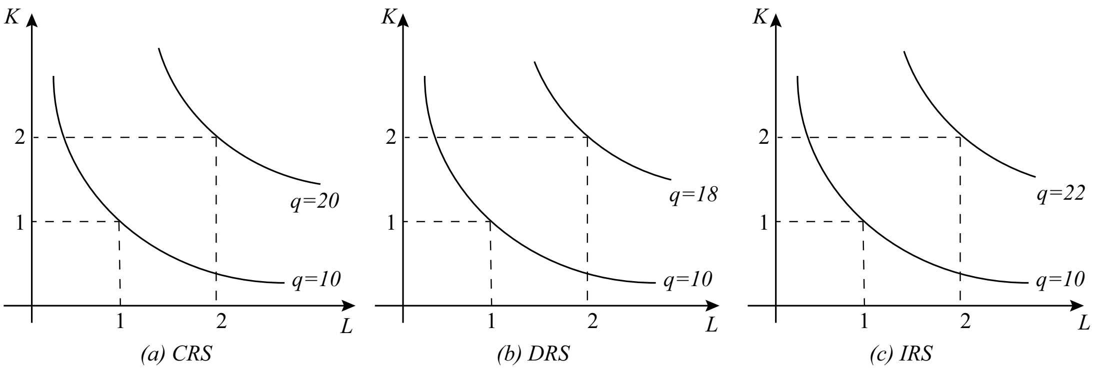

Figure 6.6 Three types of returns to scale

Returns to scale can be checked in the following way if a production function is a **homogeneous function**, i.e.,

$$F(\lambda K, \lambda L) = \lambda^t F(K, L), \qquad \lambda > 0.$$

- if the exponent $t < 1$, $F(K, L)$ displays DRS,
- if the exponent $t = 1$, $F(K, L)$ displays CRS,
- if the exponent $t > 1$, $F(K, L)$ displays IRS.

Example 6.6 Suppose

$$F(K, L) = 5K^{\frac{1}{3}}L^{\frac{2}{3}}.$$

Then

$$F(\lambda K, \lambda L) = 5(\lambda K)^{\frac{1}{3}}(\lambda L)^{\frac{2}{3}} = 5\lambda^{\frac{1}{3}+\frac{2}{3}}K^{\frac{1}{3}}L^{\frac{2}{3}} = 5\lambda K^{\frac{1}{3}}L^{\frac{2}{3}} = \lambda F(K, L).$$

Since the exponent of λ is 1, it is CRS.

Example 6.7 Suppose

$$F(K, L) = KL.$$

Then

$$F(\lambda K, \lambda L) = (\lambda K)(\lambda L) = \lambda^2 KL = \lambda^2 F(K, L).$$

Since the exponent of λ is 2, it is IRS.

For a general Cobb-Douglas production

$$F(K, L) = K^{\alpha} L^{\beta} \quad \alpha > 0,\ \beta > 0,$$

we have

$$F(\lambda K, \lambda L) = (\lambda K)^{\alpha}(\lambda L)^{\beta} = \lambda^{\alpha+\beta} K^{\alpha} L^{\beta} = \lambda^{\alpha+\beta} F(K, L).$$

Therefore, we have

- if the exponent $(\alpha + \beta) < 1$, $F(K, L)$ displays DRS,
- if the exponent $(\alpha + \beta) = 1$, $F(K, L)$ displays CRS,
- if the exponent $(\alpha + \beta) > 1$, $F(K, L)$ displays IRS.

Example 6.8 Suppose

$$F(K, L) = aK + bL.$$

Then

$$F(\lambda K, \lambda L) = a(\lambda K) + b(\lambda L) = \lambda(aK + bL) = \lambda F(K, L).$$

Since the exponent of λ is 1, it is CRS. Thus, any linear production function is CRS.

Example 6.9 Suppose the production is given by a fixed-proportion production function

$$q = F(K, L) = \min\{aK, bL\}.$$

Then

$$F(\lambda K, \lambda L) = \min\{\lambda aK, \lambda bL\} = \lambda \min\{aK, bL\} = \lambda F(K, L).$$

Since the exponent of λ is 1, it is CRS. Thus, any fixed proportional production function is CRS.

Remarks: Returns to scale need not be uniform across all possible levels of output. As a general rule, increasing returns to scale are likely prevail when the scale of operations is small, perhaps followed by an intermediate range when constant returns prevail, with decreasing returns to scale becoming more relevant for large-scale operations. In other words, a production function could embody increasing, constant, and decreasing returns to scale at different levels of output.

Chapter 7

The Cost of Production

We now discuss various types of production cost and analyze the relationship between cost of production and output of production.

7.1 The Nature of Cost

In analyzing production decision of a firm, the basic concept is the economic costs of production.

Economic costs: The ***sum*** of explicit and implicit costs.

Explicit costs (i.e., **accounting cost**): the payments explicitly made for resources which the firm purchases or hires from outside sources (such as wages, interest paid on debt, and land rent).

Implicit costs: the costs of resources which the firm uses but neither buys nor hires from outside sources.

—- Provided these resources have an ***alternative use***, there is a cost involved although no explicit monetary payment is made. They stand for the monetary payments the resources could earn in their best alternative use. For example,

- If you own a building, implicit costs of running a small store include the rent that could have been earned if the building was leased to another firm.

- Salary that could be earned by owner if employed in another business.

- Interest that could be earned by lending money to someone else.

Thus, the sum of explicit and implicit costs can be regarded as opportunity costs. That is, ***economic cost equals opportunity cost***.

Opportunity cost: Cost associated with opportunities forgone when one alternative is chosen.

—- If they are used to produce one good, they are not available for producing other goods. For example, cost of a refrigerator might be a number of washing machines.

The concept of opportunity cost is particularly useful in situations where alternatives that are forgone do not reflect monetary outlays.

Sunk cost: Expenditure that has been made and cannot be recovered (i.e., no alternative use).

Because a sunk cost cannot be recovered, it should not influence the firm's decisions. For example, if specialized equipment for a plant cannot be converted for alternative use, the expenditure on this equipment is a sunk cost. ***Because it has no alternative use, its opportunity cost is zero. Thus it should not be included as part of the firm's economic costs***.

A prospective sunk cost is an investment. Here the firm must decide whether that investment in specialized equipment is economical.

7.2 Short-Run Costs of Production

7.2.1 Short-Run Total, Average, and Marginal Costs

The short-run costs of production include:

(a) **Total Fixed Cost (TFC)**: Costs of fixed factors of production in the short-run.

Fixed cost does not vary with the level of output — it must be paid even if output equals zero.

(a') **Average Fixed Cost (AFC)**:

$$AFC = \frac{TFC}{q} \text{ where } q = \text{output level.}$$

(b) **Total Variable Cost (TVC)**: Costs incurred by the firm that depend on how much output it produces. These costs are associated with the variable inputs.

For example, labor cost depends on the number of workers hired.

(b') **Average Variable Cost (AVC)**:

$$AVC = \frac{TVC}{q}.$$

(c) **Total Cost (TC)**:

$$TC = TFC + TVC.$$

(c') **Average Total Cost (ATC)**:

$$ATC = \frac{TC}{q}.$$

Since $TC = TFC + TVC$, then

$$ATC = \frac{TFC + TVC}{q} = \frac{TFC}{q} + \frac{TVC}{q}.$$

So,

$$ATC = AFC + AVC.$$

(d) **Marginal Cost (MC)**: Increase in cost resulting from the production of one extra unit of output.

$$MC = \frac{\Delta TC}{\Delta q} = \frac{\Delta TFC + \Delta TVC}{\Delta q}.$$

Thus $\Delta TFC = 0$ because only TVC varies with q. Therefore

$$MC = \frac{\Delta TC}{\Delta q}.$$

It is noticed that $TVC =$ sum of MC. This is because fixed cost does not change as the firm's level of output changes, marginal cost is equal to the increase in variable cost or the increase in total cost that results from an extra unit of output.

Remarks:

(1) ***How do we know which costs are fixed and which are variable?*** Over a very short time horizon — say, a few months — most costs are fixed. Over such a short period, a firm is usually obligated to pay for contracted shipments of materials. Over a very long time horizon — say, ten years — nearly all costs are variable. Workers and managers can be laid off (or employment can be reduced by attrition), and much of the machinery can be sold off or not replaced as it becomes obsolete and is scrapped.

(2) ***Fixed costs versus sunk costs***. Fixed costs can be avoided if the firm shuts down or goes out of business. Sunk costs, on the other hand, are costs that have been incurred and cannot be recovered. Fixed costs affect the firm's decisions looking forward, whereas sunk costs do not. Fixed costs that are high relative to revenue and cannot be reduced might lead a firm to shut down—eliminating those fixed costs and earning zero profit may be better than incurring ongoing losses.

Incurring a high sunk cost might later turn out to be a bad decision, but the expenditure is gone and cannot be recovered by shutting down. A sunk cost is different and does affect the firm's decisions looking forward.

Example 7.1 Short-Run Cost Schedule for an Individual Firm:

q	TFC	TVC	TC	MC	AFC	AVC	ATC
0	100	0	100	$---$	$---$	$---$	$---$
1	100	90	190	90	100	90	190
2	100	170	270	80	50	85	135
3	100	240	340	70	$33\frac{1}{2}$	80	$113\frac{1}{2}$
4	100	300	400	60	25	75	100
5	100	370	470	70	20	74	94
6	100	450	550	80	$16\frac{2}{3}$	75	$91\frac{2}{3}$
7	100	540	540	90	$14\frac{2}{7}$	$77\frac{1}{7}$	$91\frac{3}{7}$
8	100	650	750	110	$12\frac{1}{2}$	$81\frac{1}{4}$	$93\frac{3}{4}$
9	100	780	880	130	$11\frac{1}{9}$	$86\frac{2}{3}$	$97\frac{7}{9}$
10	100	930	1030	150	10	93	103

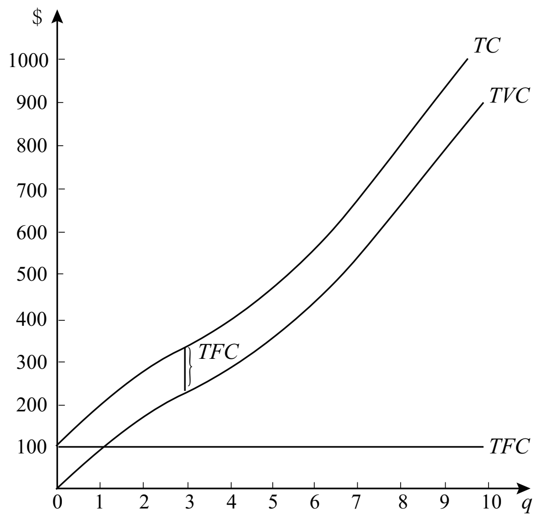

Figure 7.1 TFC is the gap between TC and TVC

Note that in Figure 7.1 the vertical distance between TC and TVC is equal to TFC. Also, in Figure 7.2, the vertical distance between AVC and ATC is AFC.

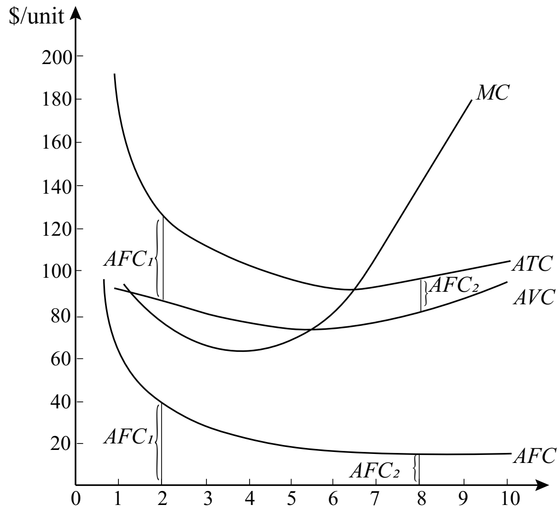

Figure 7.2 AFC is the gap between AVC and ATC

If we have a specific cost function, we can easily find the related various cost curves.

Example 7.2 Suppose that $TC = 10 + 2q + 5q^2$. Then, we have:

$$TFC = 10,\ TVC = 2q + 5q^2,\ AC = 10/q + 2 + 5q,$$

$$AFC = 10/q,\ AVC = 2 + 5q,\ MC = 2 + 10q.$$

7.2.2 The Shapes of Short-Run Cost Curves

Now let us examine what shapes the short-run cost curves have. We will show that all short-run cost curves have the same shapes as the shapes implied by the data in the previous example.

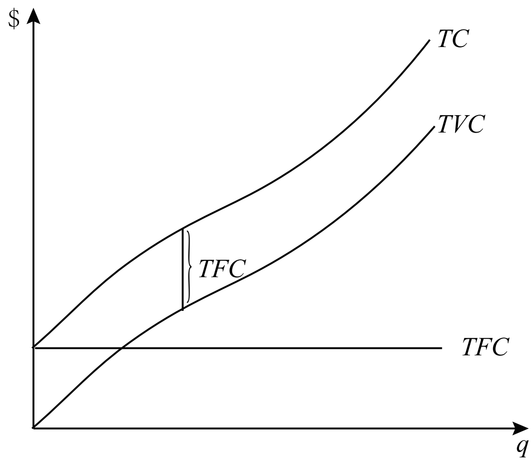

Figure 7.3 TFC is the gap between TC and TVC

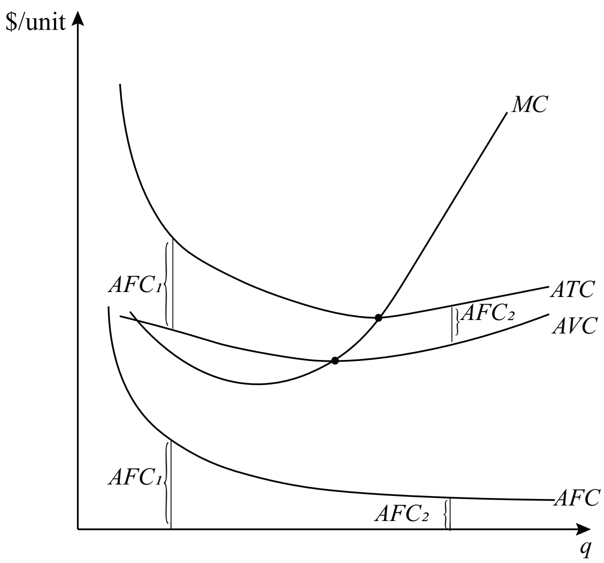

Figure 7.4 MC curve cuts curves of AVC and ATC at their respective minimum

The same as the product curves, we have:

(1) *The average total cost of a given level of output is the slope of the line from the origin to the total cost curve at that level of output*.

(2) *The marginal cost of a given level of output is the slope of the line that is tangent to the total cost curve at that level of output.*

The U-shaped Marginal Cost Curve

The marginal cost curve is generally U-shaped: With the cost of additional units of output first falling, reaching a minimum, and then rising.

The U-shape of the marginal cost curve is attributable to the law of diminishing marginal returns.

To see why, recall the MC is defined as

$$MC = \frac{\Delta TVC}{\Delta q}.$$

We know $TVC = wL$ where w is the wage rate and L is the amount of the variable input (labor). Thus $\Delta TVC = w\Delta L$, yielding

$$MC = \frac{\Delta TVC}{\Delta q} = \frac{w\Delta L}{\Delta q} = w/\frac{\Delta q}{\Delta L} = w/MP_L.$$

Thus MC ***and*** MP_L ***have a negative relationship***. Because of the law of diminishing marginal returns, MP_L varies with the amount of output and therefore, so must MC. At low levels of output MP_L is rising; so, correspondingly, $MC(= w/MP_L)$ must be falling. When MP_L reaches a maximum, then MC must be at a minimum. After that MP_L falls, and then MC must rise. That is, MP_L rises and then falls; and correspondingly MC will first fall and then rise.

The ***average variable curve (***AVC***) must be also U-shaped*** due to the law of diminishing marginal returns. Recall AVC is defined as

$$AVC = \frac{TVC}{q} = \frac{wL}{q} = w/\frac{q}{L} = w/AP_L.$$

In the previous chapter we saw that the law of diminishing marginal returns leads to an AP_L shaped like an inverted U. That is, AP_L rises, reaches a maximum, and then falls. As a result, w/MP_L (MC) must be U-shaped, i.e., MC will fall, reach a minimum, and then rise. Thus, MC ***cuts*** AVC ***and*** ATC ***at their respective minimum***.

The Average-Marginal Costs Relationship

- If $MC < AVC$ (resp. $MC < AC$), the AVC (resp. AC) must be falling.
- If $MC > AVC$ (resp. $MC > AC$), the AVC (resp. AC) must be rising.
- If $MC = AVC$ (resp. $MC = AC$), the AVC (resp. AC) must be at the minimum.

Again, when MC is constant, AVC is identical to MC.

Knowledge of short-run costs is particularly important for firms that operate in an environment in which demand conditions fluctuate considerably. If the firm is currently producing at a level of output at which marginal cost is sharply increasing, and if demand may increase in the future, management might want to expand production capacity to avoid higher costs.

7.3 Long-Run Costs of Production

In the long run:

(a) Firms can enter or exit from an industry (to be discussed later);

(b) A firm can vary **all** inputs, and thus

- the law of diminishing marginal returns does not apply;

- all costs are variable so there is no distinction between fixed and variable costs. The remaining types of cost are TC, ATC, and MC.

7.3.1 Isocost Line

Isocost line: All possible combinations of labor and capital that can be purchased for a given total cost and given prices of inputs as depicted in Figure 7.5. (Compare it with the budget line.)

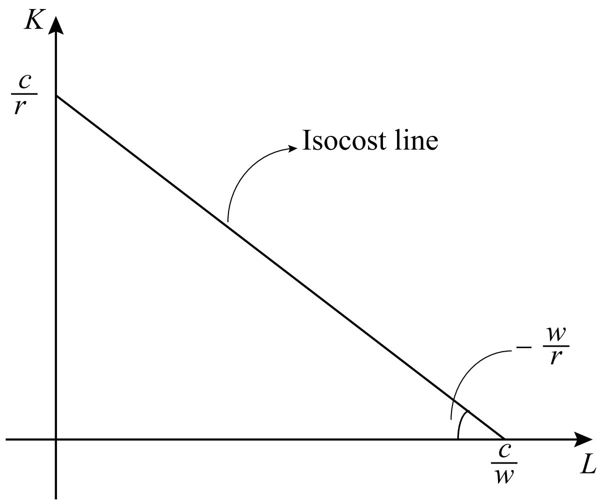

Figure 7.5 An isocost line with two inputs

To see what an isocost line looks like, recall that the total cost of producing any particular output is given by the sum of the firm's labor cost wL and its capital cost rK:

$$C = wL + rK,$$

where C is total cost, w is wage rate, r is rental rate of K, and slope of the isocost line is $-\frac{w}{r}$.

We can draw the isocost line map by varying total cost as shown in Figure 7.6.

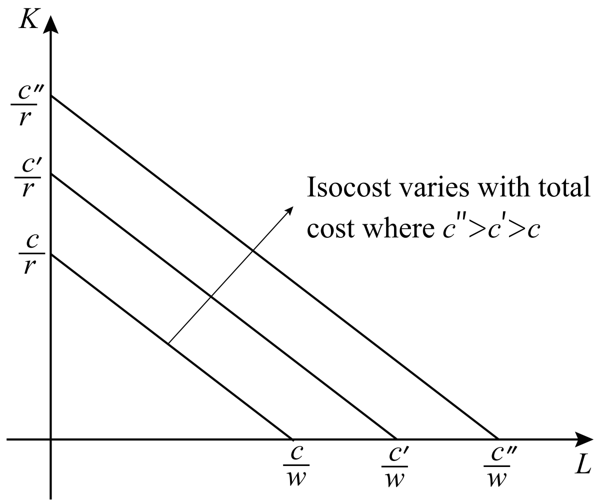

Figure 7.6 A set of isocost lines as total cost varies

Suppose that a firm wants to produce a given level of output, $q = \bar{q}$. We can use the isoquant to determine the possible input combinations which make $\bar{q}$ feasible. Which combination will be chosen?

The firm would choose whichever combination of inputs that

1) yield $\bar{q}$ units of output;
2) costs less than any other input combinations which also produce $\bar{q}$.

7.3.2 Cost Minimization Problem

The firm wishes to find the input combination which produces a given level of output and incurs the lowest possible costs. Thus, the firm must find a point on the isoquant $q = \bar{q}$ which is tangent to an isocost line.

In Figure 7.7, both D and E produce $\bar{q}$, but D costs more.

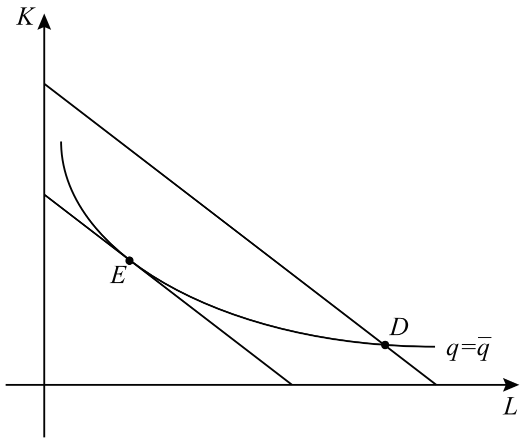

Figure 7.7 The cost-minimizing input choice

At the cost minimizing bundle, the slope of isocost generally equals the slope of isoquant, and hence we have

$$\frac{w}{r} = MRTS_{LK}. \tag{7.1}$$

Recall that we showed that the marginal rate of technical substitution of labor for capital ($MRTS_{LK}$) is the absolute value of the slope of the isoquant and is equal to the ratio of the marginal products of labor and capital: $MRTS_{LK} = \frac{MP_L}{MP_K}$, we can rewrite (7.1) as

$$\frac{MP_L}{MP_K} = \frac{w}{r}. \tag{7.2}$$

Rearranging terms, we obtain

$$\frac{MP_L}{w} = \frac{MP_K}{r}, \tag{7.3}$$

which means that the total cost of producing a given level of output is minimized when ***the ratio of marginal product to input price is equal for all inputs***. That is, the firm should employ inputs such that the marginal product cost per dollar's worth of all inputs is equal.

Expansion path: Line which shows input combinations that lead to cost minimization, and it is formed by connecting all tangency points when the level of output varies.

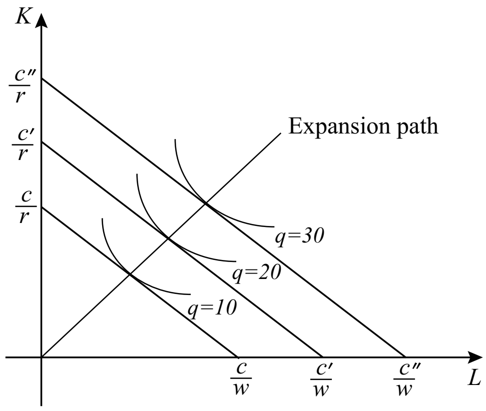

Figure 7.8 The expansion path

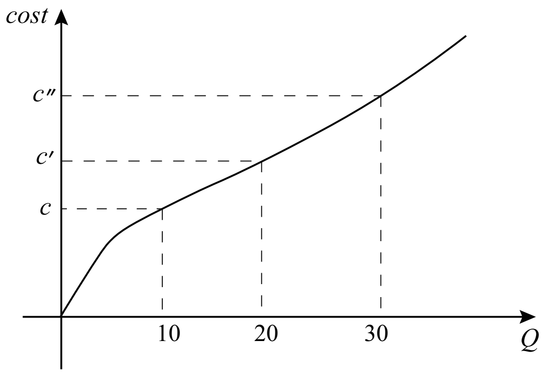

Figure 7.9 Generating a total cost curve through expansion path

We can use the expansion path to generate a total cost curve. To move from the expansion path to the cost curve, we follow three steps:

- Choose an output level represented by an isoquant. Then find the point of tangency of that isoquant with an isocost line.

- From the chosen isocost line, determine the minimum cost of producing the output level that has been selected.

- Graph the output-cost combination in Figure 7.9.

7.3.3 Long-Run Cost Curves

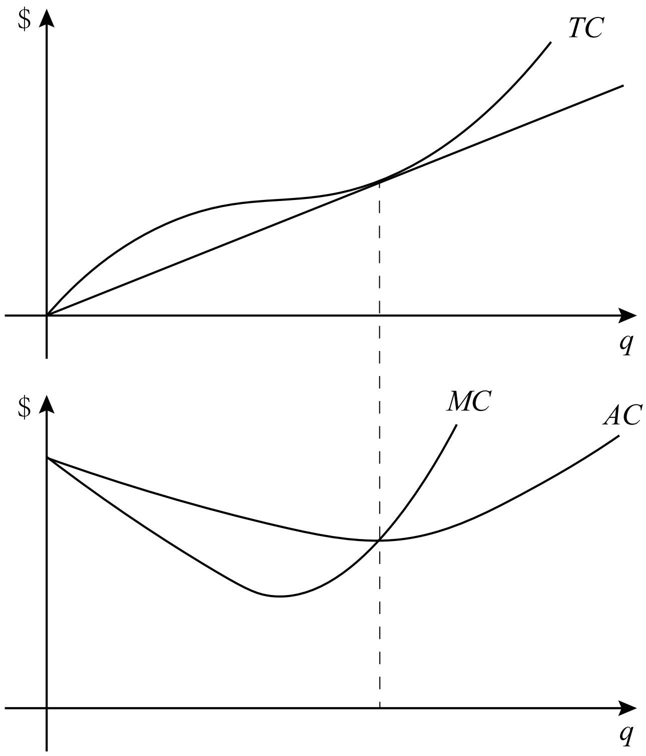

Figure 7.10 The relationship between TC, MC and AC

The long-run total cost shows the minimum cost at each level of output. The long-run marginal cost and average cost curves are derived from the long-run total cost curve in the same way that the short-run per-unit curves are derived from the short-run total cost curves.

We have drawn the long-run TC curve which yields a U-shaped long-run AC and MC curves. Why would the AC and MC curves have this shape? In the long-run all inputs are variable, and hence the law of diminishing marginal returns is not responsible for their U-shapes.

The shape of the long-run AC and MC curves reflects the feature of returns to scale of the production technology. As we explained before, increasing returns to scale are likely to be common at low rates of output, while decreasing returns to scale are likely to prevail at high output levels. Therefore, the long-run AV and MC curves must have a U-shape.

Increasing Returns to Scale

Double input cost ⇒ more than double the output.

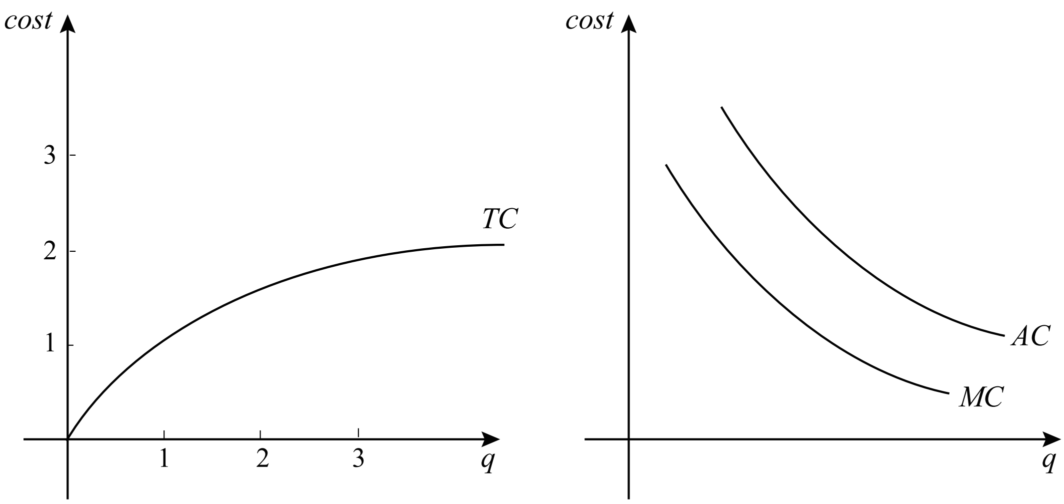

Figure 7.11 Cost curves under increasing returns to scale

MC = slope of TC falls, and AC = slope of connecting line also falls.

Decreasing Returns to Scale

Double input cost ⇒ less than double the output. MC increases, and AC increases.

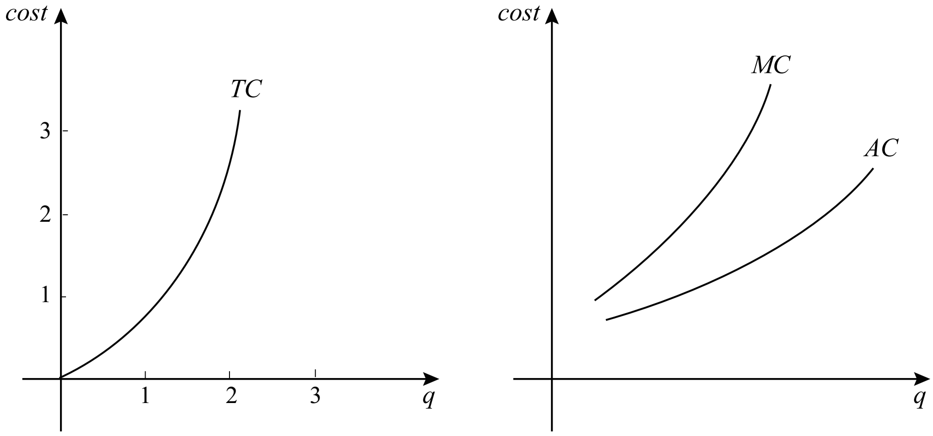

Figure 7.12 Cost curves under decreasing returns to scale

Constant Returns to Scale

Double input $\Rightarrow$ double output. Thus, input cost per unit is constant, i.e., $TC = aq$ where a is cost of producing one unit of output. Then we have:

$$AC = \frac{TC}{q} = \frac{aq}{q} = a,$$

$$MC = \frac{\Delta TC}{\Delta q} = \frac{a\Delta q}{\Delta q} = a.$$

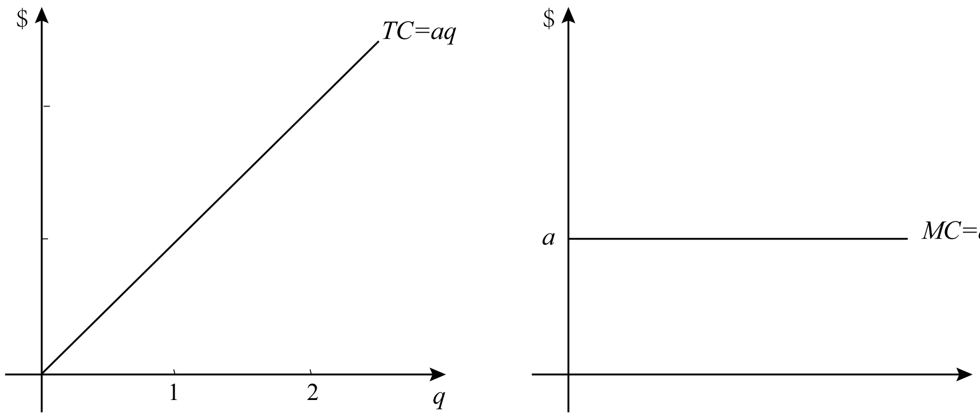

Figure 7.13 Cost curves under constant returns to scale

Note: MC curve always passes through the lowest point of the AC curve. When AC is falling, MC is below it. When AC is rising, MC is above it.

7.3.4 Input Price Changes and Cost Curves

A change in input price will cause the entire cost curves to shift.

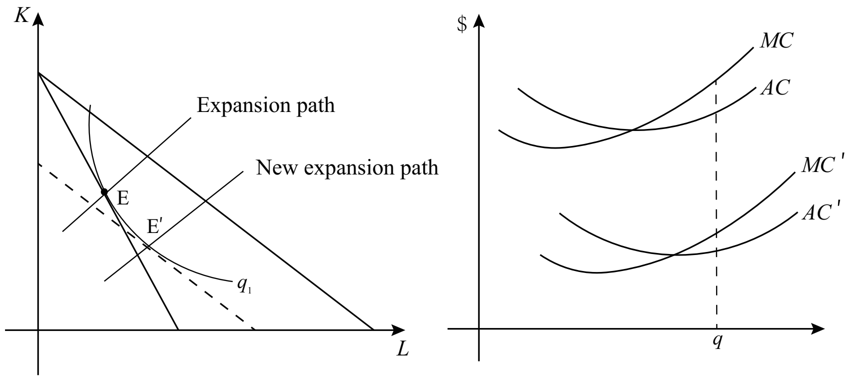

Figure 7.14 The effect of input price changes on cost curves

Initially, the firm is producing q_1, by employing K and L at point E. With a lower wage rate the cost of producing each level of output falls. Input combination E' becomes the least costly way to produce the same q_1 after the wage reduction. Thus the change in w shifts the AC and MC curves downward to AC' and MC'.

7.3.5 Economies of Scale and Diseconomies of Scale

More general and closely related concepts to returns to scale are economies of scale and diseconomies of scale.

Economies of scale: Situation in which output can be doubled for less than a doubling of cost.

Diseconomies of scale: Situation in which a doubling of output requires more than a doubling of cost.

A firm enjoys economies of scale when it can double its output for less than twice the cost. The term economies of scale includes increasing returns to scale as a special case, but it is more general because it reflects input proportions that change as the firm changes its level of production.

Increasing Returns to Scale: Output more than doubles when the quantities of all input are doubled.

Economies of scale are often measured in terms of a **cost-output elasticity**, denoted as E_C. E_C is the percentage change in the cost of production resulting from a 1-percent increase in output:

$$E_c = \frac{\Delta C/C}{\Delta q/q}.$$

To see how E_C relates to our traditional measures of cost, rewrite equation as follows:

$$E_c = \frac{\Delta C/\Delta q}{C/q} = \frac{MC}{AC}.$$

Clearly, E_C is equal to 1 when marginal and average costs are equal. In that case, costs increase proportionately with output, and there are neither economies nor diseconomies of scale. When there are economies of scale (when costs increase less than proportionately with output), marginal cost is less than average cost and E_C is less than 1. Finally, when there are diseconomies of scale, marginal cost is greater than average cost and E_C is greater than 1.

Thus, analogue to the analysis of return to scales, because of economies of and then diseconomies of scale with output increasing, the long-run AV and MC curves must have a U-shape.

7.4 Short-Run Verses Long-Run Average Cost Curves

(1) **Short-Run** AC **Curve**:

Consider five different plant capacities.

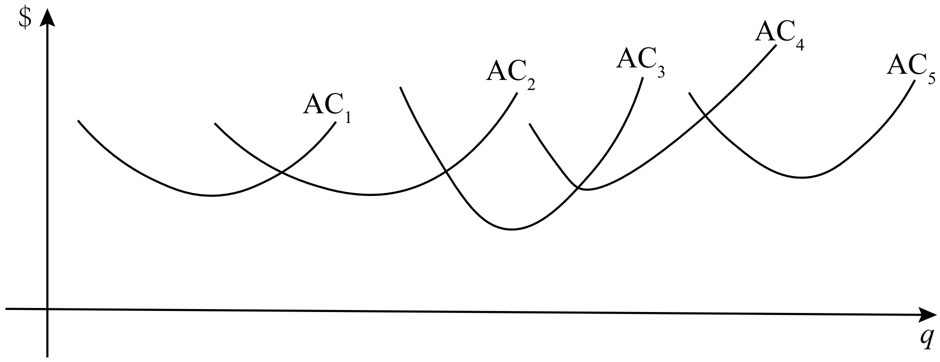

Figure 7.15 AC curves under five different plant capacities

- Plant capacities 1 & 2: small firms;
- Plant capacities 3: medium firms;
- Plant capacities 4 & 5: large firms.

Observe that the short-run ATC declines from small to medium sized firms and then increases as firms become large.

(2) **Long-Run** AC **Curve**:

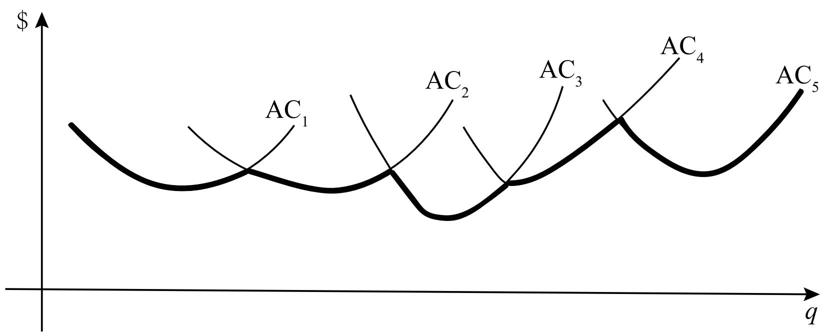

Figure 7.16 Generating the long-run AC curve

The long-run AC curve shows the lowest per unit cost at which any output can be produced, given that the firm has sufficient time to vary all inputs, including plant capacity.

The long-run AC curve consists of segments of the short-run AC curves. Given an unlimited number of possible capacity levels of a plant the long-run AC curve is made up of points of tangency with an unlimited number of short-run AC curves. Therefore, the long-run AC curve shows the most efficient way to produce a given level of output.

7.5 Economies and Diseconomies of Scope

In many situations, a company can take the cost advantages of providing a variety of products rather than specializing in the production of a single product because of economies of scope.

Product transformation curve: Curve showing the various combinations of two different outputs (products) that can be produced with a given set of inputs.

The product transformation curves O_1 and O_2 in Figure 7.17 are bowed out (i.e., concave) because there are economies of scope in production.

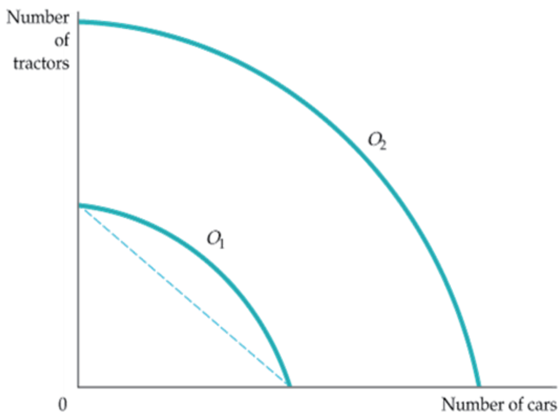

Figure 7.17 product transformation curve

Economies of scope: Situation in which joint output of a single firm is greater than output that could be achieved by two different firms when each produces a single product.

In other words, the total cost of joint production is lower than that they are produced separately.

Diseconomies of scope: Situation in which joint output of a single firm is less than could be achieved by separate firms when each produces a single product.

The Degree of Economies of Scope

To measure the degree to which there are economies of scope, we should ask what percentage of the cost of production is saved when two (or more) products are produced jointly rather than individually.

Degree of economies of scope (SC): Percentage of cost savings resulting when two or more products are produced jointly rather than individually:

$$SC = \frac{C(q_1) + C(q_2) - C(q_1, q_2)}{C(q_1, q_2)}.$$

7.6 Using Cost Curves: Controlling Pollution

Many problems can be clarified by posing them in terms of marginal cost. Here is an example of using cost curves to controlling pollution in the cheapest way. There are two firms which release pollutants into the air in the process of production. The government steps in and restricts the total pollution to a certain level, say 200 units.

In Figure 7.15, the amount of pollution generated by each firm is measured from right to left. For example, before the government restricts its activity, firm A discharges OP_1 (300 units), and firm B discharges OP_2 (250 units). Measuring pollution from right to left is the same as measuring pollution abatement - the number of units by which pollution is reduced from its initial level - from left to right. For example, if firm B cuts back its pollution from 250 to 100 units, it has produced 150 units of pollution abatement, the distance p_2X.

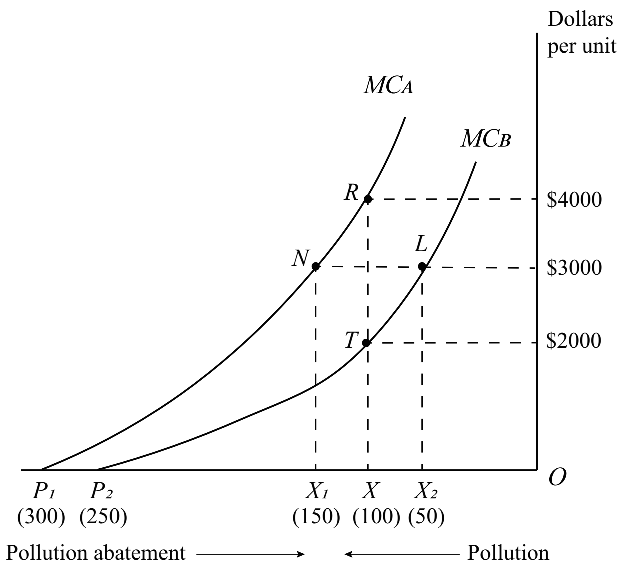

Figure 7.18 Using cost curves to illustrate pollution controls

One way to reduce pollution to 200 is to ask each firm reduces to 100, this way may not be the cheapest way to do it. At 100 units of pollution, the marginal cost of reducing pollution to firm A is $4,000, but firm B's cost only $2,000. Thus, if firm B reduces one more unit of pollution, it would add only $2,000. But if we let firm A increase one more unit of pollution, its cost would fall by $4,000. As a result, we can reduce the two firms' combined cost by $2,000.

In fact, as long as the marginal costs differ, the total cost of pollution abatement can be reduced by increasing abatement where its marginal cost is less and reducing abatement where its marginal cost is higher. Thus, to minimize the cost of pollution control, firms should produce at a point where their marginal costs are equal. To reduce pollution to 200 units in the cheapest way, firm A should discharges 150 units and firm B, 50 units.

Chapter 8

Profit Maximization and Competitive Firm

A **competitive firm** has the following two basic characteristics:

(a) The firm is one of many which produce identical products.
(b) The firm is a price taker in both input and output markets, i.e., it lacks enough market power to dictate the prices in a market. This means that the corresponding demand curve the firm faces is horizontal.

8.1 Demand Curve under Competition

Figure 8.1 shows the demand curve faced by a competitive firm. A competitive firm supplies only a small portion of the total output of all the firms in an industry. Therefore, the firm takes the market price of the product as given, choosing its output on the assumption that the price will be unaffected by the output choice. In the right graph of Figure 8.1 the demand curve facing the firm is **perfectly elastic**, even though the market demand curve in the left graph of Figure 8.1 is downward sloping.

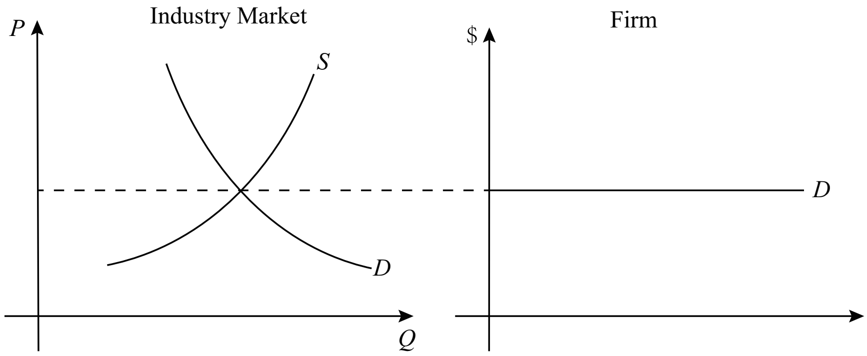

Figure 8.1 Demand curve under perfect competition

Economic profit takes into account opportunity costs. One such opportunity cost is the return to the firm's owners if their capital were used elsewhere.

Accounting profit equals revenues R minus labor cost wL, which is positive. Economic profit π, however, equals revenues R minus labor cost wL minus the capital cost, rK.

Thus we assume that the firm's objective is to maximize profit (π) where

$$\text{Profit} = \text{Total Revenue} - \text{Total Cost},$$

$$\pi = TR - TC.$$

The assumption of profit maximization predicts business behavior reasonably accurately and avoids unnecessary analytical complications. For smaller firms, profit is likely to dominate almost all other decisions. In larger firms, however, managers who make day-to-day decisions usually have little contact with the owners. Managers may be more concerned with such goals as revenue maximization, revenue growth.

Firms that do not come close to maximizing profit are unlikely to survive. The firms that do survive make long-run profit maximization one of their highest priorities.

Total revenue (TR), average revenue (AR), and marginal revenue (MR)

for a competitive firm are given by

$$TR = P \times q; \tag{8.1}$$
$$AR = \frac{TR}{q} = \frac{P \times q}{q} = P; \tag{8.2}$$
$$MR = \frac{\Delta TR}{\Delta q} = \frac{P\Delta q}{\Delta q} = P, \tag{8.3}$$

receptively.

For a competitive firm, we thus have:

$$P = AR = MR.$$

Therefore, ***along this demand curve, marginal revenue, average revenue, and price are all equal***.

8.2 Short-Run Profit Maximization

A competitive firm has no control over prices, so a profit-maximizing policy must be related to the quantity produced.

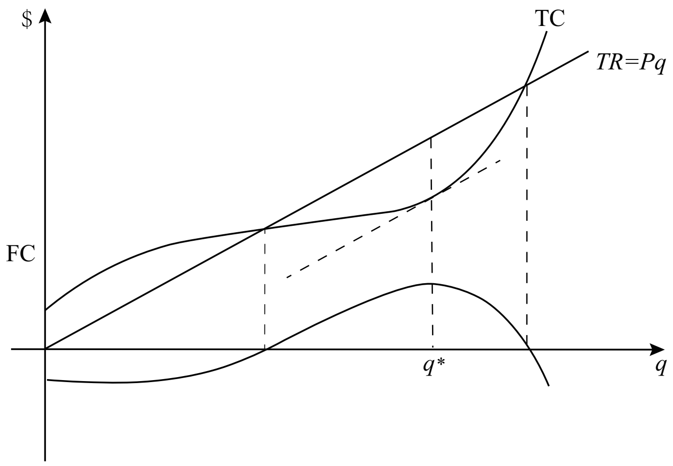

Figure 8.2 Identifying the most profitable output level using TR and TC curves

Figure 8.2 shows how we identify the most profitable level of output by using the typical TR and TC curves.

Note that π is the distance between TR and TC, and is greatest at q^*, and also that at q^* the slopes of the TR and TC curves are equal.

Should the firm produce any output?

(i) Yes, if there is a level of output which can earn a (positive) profit, i.e., if $TR > TC$ (i.e., $AR > AC$).

(ii) Yes, even if it cannot make a profit but it can make a loss smaller than fixed cost, i.e., if $TR > TVC$ (i.e., $AR > AVC$). Why?

- if $q = 0$, $TC = TFC$; $TR = 0$, so profit $= TR - TC = -TFC \Rightarrow$
$$\text{Loss} = TFC;$$
- if $q > 0$ and loss occurs, then
$$\begin{aligned}\text{Loss} &= TC - TR \\ &= (TFC + TVC) - TR \\ &= TFC + (TVC - TR)\end{aligned}$$
Therefore, Loss $< TFC$ provided $TVC < TR$ (or equivalently $AVC < AR$).

In (ii) above it is important to recall that in the short run a firm must pay its TFC regardless of the level of output. By producing output ($q > 0$) the firm may be able to reduce its loss from the no-production case ($q = 0$).

Therefore, ***the firm should produce a positive level of output as long as*** $TR > TVC$ ***(or equivalently*** $AR > AVC$***)***. Since $P = AR = MR$, we have

$$P > AVC.$$

What Quantity Should the Firm Produce when $P > AVC$?

The firm should produce the quantity which maximizes profits or minimizes losses:

- It is worthwhile producing for which $MR > MC$.
- It is not worthwhile continuing to produce when $MR < MC$. (Recall that MC must rise eventually).
- Thus a firm will produce up to the level where $MC = MR$, or $MC = P$ since $P = MR$. That is, at that output, marginal revenue (the slope of the revenue curve) is equal to marginal cost (the slope of the cost curve).

Summarizing the above discussions, we finally reach the following conclusion.

Conditions for Short-Run Profit Maximization:

1. $P \geq AVC(q^*)$ (necessary condition for producing);
2. $P = MC(q^*)$ (how much should be produced).

Otherwise, a competitive firm should shut down (i.e., $q^* = 0$) if price is below AVC.

Remarks:

(1) If condition 1 were not satisfied, then the firm is best off producing $q^* = 0$ units so that $TR = 0$, $TC = TFC + 0 = TFC$ and thus $\pi = -TC$ (losing TFC is better than losing TFC + some VC).

(2) When $P = MR(q^*) = AVC(q^*)$ and condition 2 is satisfied, then both q^* and $q = 0$ generate the same π (i.e., $\pi = -TFC$).

Various Situations in Short-Run Profit Maximization

There are five cases for a competitive firm's profit maximization, depending on relationships among P, $AVC(q^*)$, and $AC(q^*)$:

Case 1: $P > AC(q^*)$, resulting in positive profit;

Case 2: $P = AC(q^*)$, zero profit;

Case 3: $AC(q^*) > P > AVC(q^*)$, loss minimization;

Case 4: $AC(q^*) > P = AVC(q^*)$, no difference between producing $q*$ and shutdown ($q = 0$) since Loss = TFC in either case;

Case 5: $P < AC(q^*)$, shutdown.

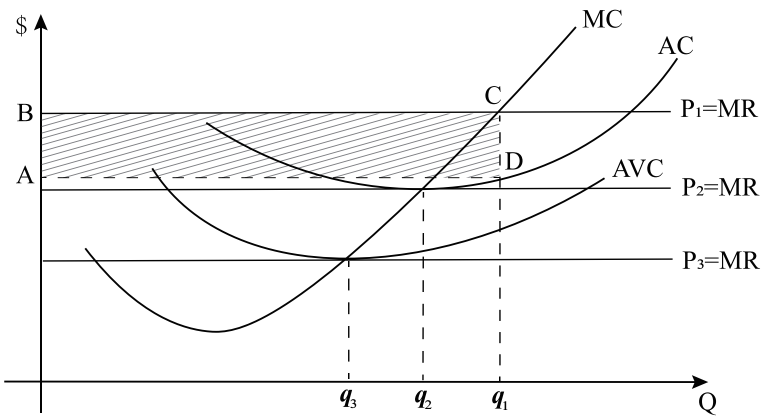

Figure 8.3 Profit maximization in the short run

As shown in Figure 8.3, when the price is P_1, the profit-maximizing output is q_1. The total profits are the rectangle $ABCD$. When the price is P_2, the most profitable output is q_2. The profits, however, are zero since the price just equals average cost of production. When the price is P_3, the firm can just cover its variable cost by producing q_3, where AVC equals the price. At this point the firm would be operating at a short-run loss; the loss is exactly equal to its total fixed cost. At a price below P_3 the firm is unable to cover its variable cost and hence must shut down.

Let us consider several numerical examples.

Example 8.1 (Perfect Competition: Profit Maximization) Note that $TC = ATC \times q$, and $TR = p \times q$. From Table 8.1, we know $MR > MC$ for $q = 1, 2, ..., 9$. $MR < MC$ for $q = 10$. $q = 9$ for profit maximization. Since $AR > ATC$ at $q = 9$, a positive profit is made.

q	AFC	AVC	ATC	MC	TC	$P = AR = MR$	TR	π
0	$---$	$---$	$---$	$---$	100	131	0	100
1	100	90	190	90	190	131	131	-59
2	50	85	135	80	270	131	262	-8
3	$33\frac{1}{3}$	80	$113\frac{1}{3}$	70	340	131	393	53
4	25	75	100	60	400	131	524	124
5	20	74	94	70	470	131	655	185
6	$16\frac{2}{3}$	75	$91\frac{2}{3}$	80	550	131	786	236
7	$14\frac{2}{7}$	$77\frac{1}{7}$	$91\frac{3}{7}$	90	640	131	917	277
8	$12\frac{1}{2}$	$81\frac{1}{4}$	$93\frac{3}{4}$	110	750	131	1048	298
9	$11\frac{1}{9}$	$86\frac{2}{3}$	$97\frac{7}{9}$	130	880	131	1179	299
10	10	93	103	150	1030	131	1310	280

Table 8.1

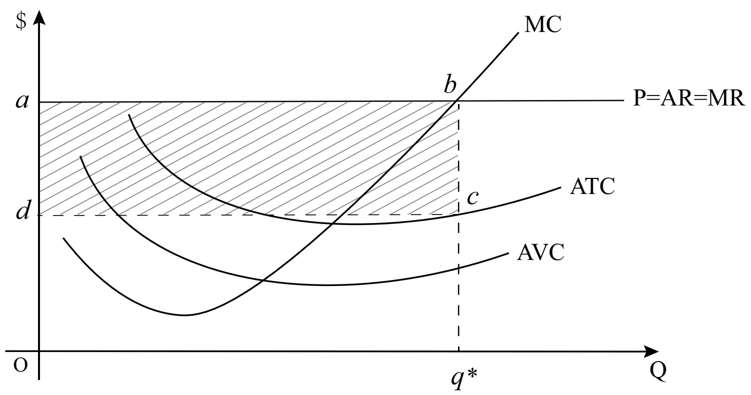

Figure 8.4 Identifying profit-maximizing output and short-run profit

From Figure 8.4, we know $TR = AR \times q^* = Oabq^*$, $TC = ATC \times q^* = Odcq^*$, Profit $= TR - TC = abcd$. At point b in Figure 8.4, $MR = MC \Rightarrow q^* = 9$.

Example 8.2 (Perfect Competition: Loss Minimization) By Table 8.2, $MR = MC$ at $q = 6$. Also, $ATC > AR > AVC$ implies loss minimization. Indeed, At $q = 6$, Loss $= \$64 < TFC = \$100$, so firm should not cease production.

q	AFC	AVC	ATC	MC	TC	$P = AR = MR$	TR	π
0	$---$	$---$	$---$	$---$	100	81	0	-100
1	100	90	190	90	190	81	81	-109
2	50	85	135	80	270	81	162	-108
3	$33\frac{1}{3}$	80	$113\frac{1}{3}$	70	340	81	243	-97
4	25	75	100	60	400	81	324	-76
5	20	74	94	70	470	81	405	-65
6	$16\frac{2}{3}$	75	$91\frac{2}{3}$	80	550	81	486	-64
7	$14\frac{2}{7}$	$77\frac{1}{7}$	$91\frac{3}{7}$	90	640	81	517	-73
8	$12\frac{1}{2}$	$81\frac{1}{4}$	$93\frac{3}{4}$	110	750	81	567	-102
9	$11\frac{1}{9}$	$86\frac{2}{3}$	$97\frac{7}{9}$	130	880	81	729	-151
10	10	93	103	150	1030	81	810	-220

Table 8.2

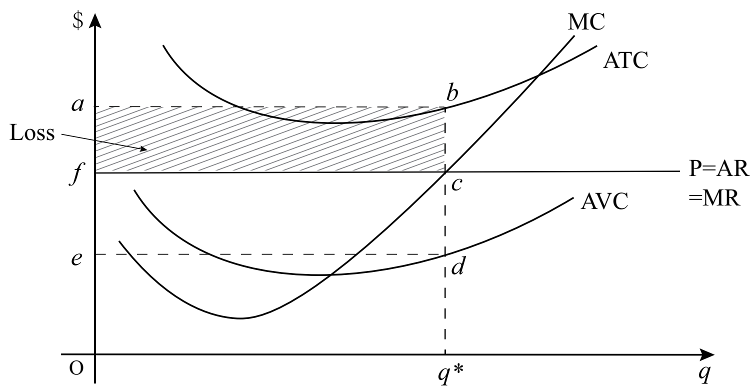

Figure 8.5 Loss minimization under perfect competition

From Figure 8.5, we know $TR = AR \times q^* = Ofcq^*$, $TC = Oabq^*$, and Loss $= TC - TR = abcf$. At point c in Figure 8.5, $MR = MC \Rightarrow q^* = 6$.

Example 8.3 (Perfect Competition: Production Shutdown) From Table 8.3, although $MR = MC$ at $q = 5$, but $AR < AVC$ at $q = 5$, which implies Loss $> TFC$ $(115 > 100)$. Thus $q = 0$, i.e., shutdown is optimal.

q	AFC	AVC	ATC	MC	TC	$P = AR = MR$	TR	π
0	$---$	$---$	$---$	$---$	100	$---$	$---$	-100
1	100	90	190	90	190	71	71	-119
2	50	85	135	80	270	71	142	-128
3	$33\frac{1}{3}$	80	$113\frac{1}{3}$	70	340	71	213	-127
4	25	75	100	60	400	71	284	-116
5	20	74	94	70	470	71	355	-115
6	$16\frac{2}{3}$	75	$91\frac{2}{3}$	80	550	71	426	-124
7	$14\frac{2}{7}$	$77\frac{1}{7}$	$91\frac{3}{7}$	90	640	71	497	-143
8	$12\frac{1}{2}$	$81\frac{1}{4}$	$93\frac{3}{4}$	110	750	71	568	-182
9	$11\frac{1}{9}$	$86\frac{2}{3}$	$97\frac{7}{9}$	130	880	71	639	-241
10	10	93	103	150	1030	71	710	-320

Table 8.3

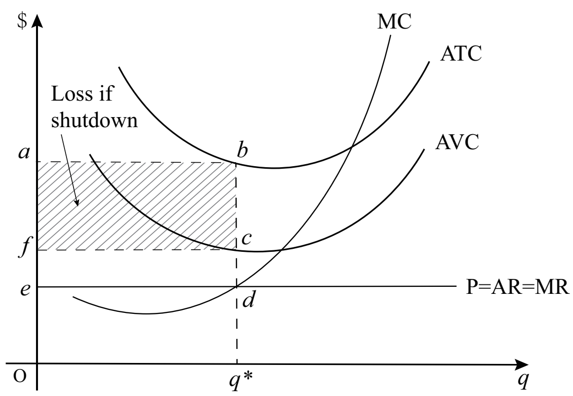

Figure 8.6 Production shutdown under perfect competition

By Figure 8.6, we know at point d in Figure 8.6, $MR = MC \Rightarrow q^* = 5$. At $q^* = 5$, $TR = Oedq^*$, $TC = Oabq^*$, $\Rightarrow$ Loss $= eabd$. At $q = 0$, $TR = 0$, $TC = TFC = AFC \times q = abcf$, which is small than $eabd$.

The following is an algebraic example.

Example 8.4 Suppose $TC(q) = 200 + 9q + 5q^2$, $MC(q) = 9 + 10q$, and $P = 16$.

(a) Find the AVC. $AVC = \frac{VC}{q}$ where VC is the portion of TC which involves terms with output q:

$$VC(q) = 9q + 5q^2.$$

Thus $AVC = \frac{9q+5q^2}{q} = 9 + 5q$.

(b) Find the profit-maximizing level of output.

- $P = MC(q^*) \Rightarrow 16 = 9 + 10q \Rightarrow q^* = \frac{7}{10}$.
- $P \geq AVC(q^*)$?
  $AVC(q^*) = 9 + 5q^* = 9 + 5 \times \frac{7}{10} = 9 + \frac{7}{2} = 12.5 < 16$.

Thus $P \geq AVC(q^*)$, and hence $q^* = \frac{7}{10}$ is indeed profit-maximizing level of output.

Note: If we change P, we find a new q^*. Repeating this we can draw a supply curve.

8.3 Competitive Firm's Short-run Supply Curve

From the conditions for short-run profit maximization, we know that the firm chooses its output so that marginal cost MC is equal to price as long as it can cover its average variable cost, and also $MC = P$ shows that a competitive firm will produce more at higher price levels because increased production becomes profitable at higher prices.

Thus, the ***firm's supply curve is the portion of the marginal cost curve that lies above the minimum average variable cost***.

The short-run supply curve is given by the crosshatched portion of the marginal cost curve as in Figure 8.7.

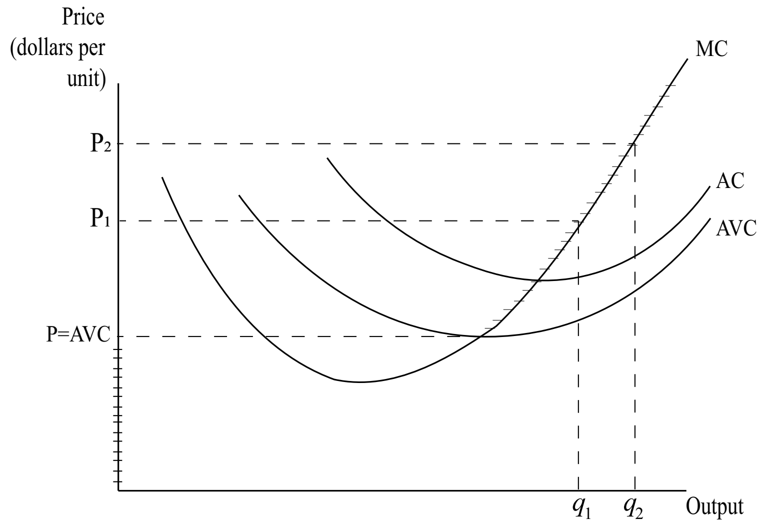

Figure 8.7 Output response to an input price reduction

8.4 Output Response to a Change in Input Prices

A change in the price of an input, combined with an unchanged output price, will change the profit-maximizing level of output. If the price of an input falls, MC shifts to MC', and output increases, where MC' equals the unchanged output price.

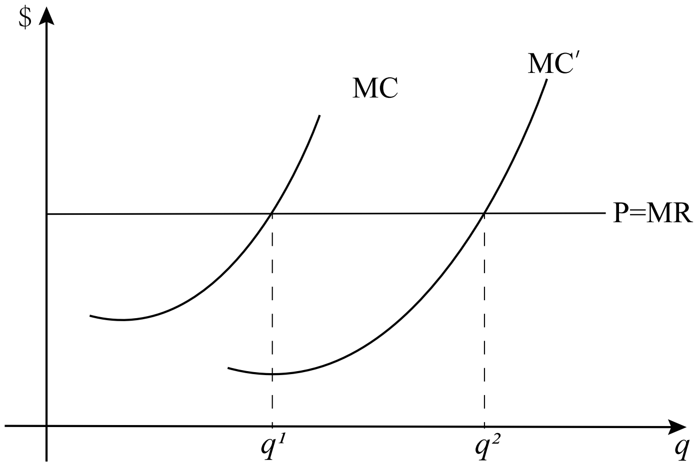

Figure 8.8 Output response to an input price reduction

8.5 Producer Surplus in the Short Run

Producer surplus: Sum of differences between the market price of a good and the marginal cost for all units produced.

Producer Surplus for a Firm

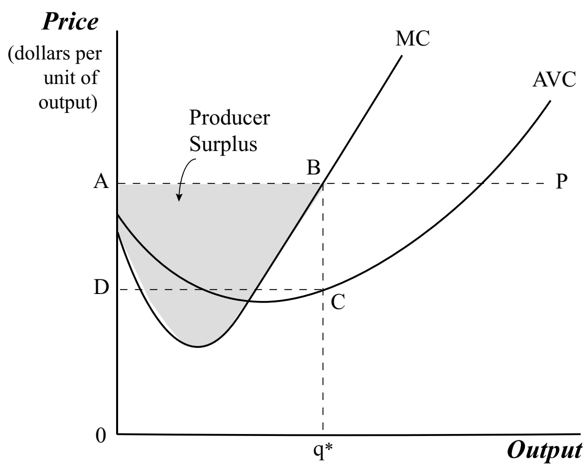

Figure 8.9 Producer surplus for a firm

The producer surplus (PS) for a firm is measured by the shaded area below the market price and above the marginal cost curve, between outputs 0 and q*, the profit-maximizing output in Figure 8.9

Alternatively, it is equal to rectangle $ABCD$ because the sum of all marginal costs up to q^* is equal to the variable costs of producing q^*. That is,

$$PS = TR - TVC.$$

The difference between producer surplus and profit of a firm is then the total fixed cost since

$$\pi = TR - TVC - TFC = PS - TFC.$$

Producer Surplus for a Market

The producer surplus for a market is the area below the market price and above the market supply curve, between 0 and output Q* as depicted in Figure 8.10.

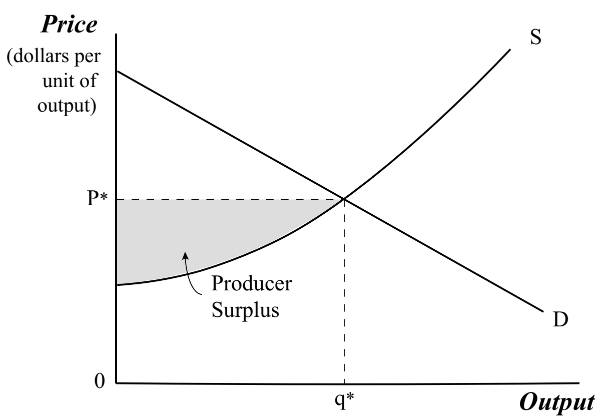

Figure 8.10 Producer surplus for a market

8.6 Long-Run Profit Maximization

The same principle used for the short-run setting can apply to the discussion of long-run profit maximization, but now we employ long-run cost curves. The firm maximizes profits in the long run by producing up to where $P = MC$. That is, the long-run output of a profit-maximizing competitive firm is the point at which long-run marginal cost equals the price.

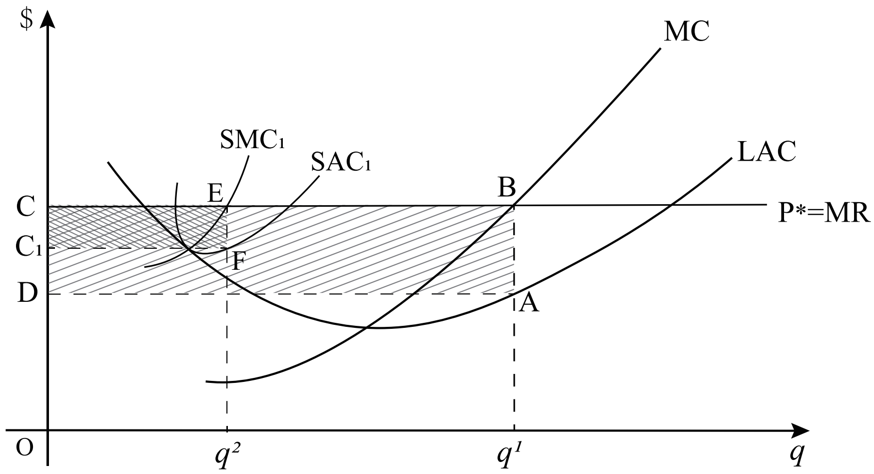

Figure 8.11 The long-run output of a profit-maximizing competitive firm

With a price P^*, the most profitable output will be q^1 in the long run, and the profit is rectangle $ABCD$. In the short run the most profitable output will be q^2, and rectangle $CEFC_1$ is the profit which is less than the long-run profit.

Conditions for Long-Run Profit Maximization:

(1): $P \geq AC$ (necessary condition for producing);

(2): $P = MC$ (how much should be produced).

Note that producer surplus coincides with profit of a firm since the total fixed cost equals zero in the long-run.

Chapter 9

Competitive Markets

In this chapter the emphasis shifts from the individual firm to the competitive industry/markets.

Assumptions on Perfect Competitive Markets:

- **A large number of buyers and sellers**, which will normally guarantee that the firms and consumers behave as price takers.
- **Unrestricted mobility of resources**: no barrier to entry into, or exit from the market.
- **Homogeneous products**: the products of all of the firms in an industry are identical to consumers.
- **Possession of all relevant information** to make economic decisions.

Unrestricted mobility implies that there are no special costs that make it difficult for a firm to enter (or exit) an industry. With free entry and exit, buyers can easily switch from one supplier to another, and suppliers can easily enter or exit a market.

The presence of many firms is not sufficient for an industry to approximate perfect competition since firms can implicitly or explicitly collude in setting prices. Conversely, the presence of only a few firms in a market does not rule out competitive behavior.

When Is a Market Highly Competitive? Many markets are highly competitive in the sense that firms face highly elastic demand curves and relatively easy entry and exit, but there is no simple rule of thumb to describe whether a market is close to being perfectly competitive.

9.1 The Short-Run Industry Supply Curve

In the short run a competitive firm will produce at a point where the marginal cost equals the price, as long as the price is above the minimum point of its average variable cost curve.

How is the short-run industry supply curve determined?

The horizontal sum of MC curves for each firm (above AVC) gives the industry supply curve. That is, the short run industry supply curve is derived by simply adding the quantities produced by all individual firms.

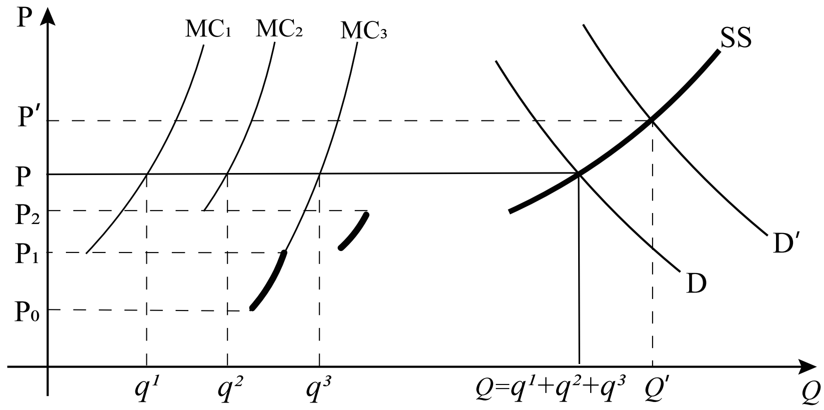

Figure 9.1 The short-run industry supply curve

Note that the short-run supply curve, SS, slopes upward. Recall that each firm's marginal cost curve slopes upward because it reflects the law of diminishing marginal returns to variable inputs. Thus, the law of diminishing marginal returns is the underlying determinant of the shape of a competitive industry's short-run supply curve.

Price and Output Determination in the Short Run

The interaction of supply and demand in a market determines market price and output. In Figure 9.1 the intersection of the demand curve D with the supply curve SS pins down the price where total quantity demanded equals total quantity supplied. Thus P is the equilibrium price and equilibrium industry output is Q, where $Q = q^1 + q^2 + q^3$.

In the short run an increase in market demand leads to a higher price and higher output. When demand increases to D', the equilibrium price becomes P' and industry output is now Q'.

9.2 Long-Run Competitive Equilibrium

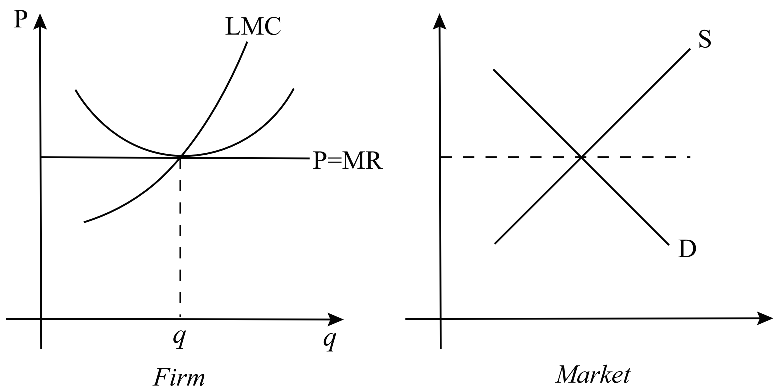

Figure 9.2 The emergence of long-run competitive equilibrium

In the long-run competitive equilibrium, the independent plans of firms and consumers mesh perfectly. Each firm has adjusted its scale of operation in light of the prevailing price and is able to sell as mush as it chooses to sell. Consumers are able to purchase as much as they wish to consume at the prevailing price. There are no incentives for any firm to alter its scale of operation or to exit from the industry and no incentive for outsiders to enter the market. Unless the underlying market conditions change, the price and rate of output will remain stable.

Long-Run Competitive Equilibrium:

(a) **Profit maximization**: Each firm must be producing the output level such that its profit is maximized at the prevailing market price, that is, $P = LMC$.

(b) **Zero economic profit made by each firm**. There are no incentives for firms to either enter or exit from the industry.

(c) **Market quantity supplied equals market quantity demanded**: The combined quantity of outputs of all firms at the prevailing price must equal the total quantity consumers wish to purchase at that price level.

Zero economic profit: A firm is earning a normal return on its investment in a perfectly competitive market — i.e., it is doing as well as it could by investing its money elsewhere.

Remarks: There are instances in which firms earning positive accounting profit may be earning zero economic profit. If a clothing store is be located near a large shopping center, the additional flow of customers can increase the store's accounting profit. When the opportunity cost of the land is included, the profitability of the clothing store is no higher than that of its competitors.

Entry and exit

In a market with entry and exit, a firm enters when it can earn a positive long-run profit and exits when it faces the prospect of a long-run loss.

Entry of Firms Eliminates Profits:

(i) In Figure 9.3, the representative firm is in a long-run equilibrium. P^e = minimum point of ATC. Normal profits are earned, and hence economic profit $= 0$.

(ii) Suppose demand increases from D^1 to D^2. Price rises to p_1 and exceeds the minimum point of ATC. Thus, positive economic profits are earned.

(iii) New firms enter the industry. Supply starts to increase from S_1, and then price starts to fall.

(iv) Supply continues to increase while economic profits are being made. The increase of supply will not stop until it reaches the new supply curve S_2, where P^e = minimum point of ATC again.

(v) In the long run, a larger quantity is supplied at the same initial price level.

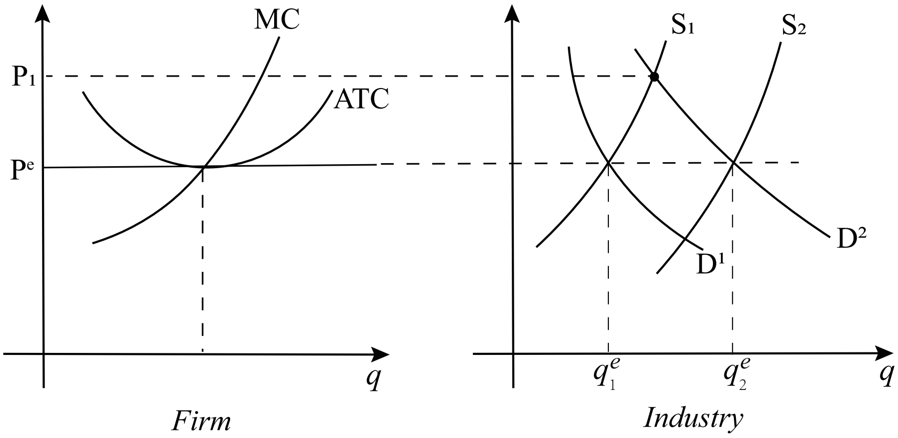

Figure 9.3 Entry of firms eliminates economic profits

Exit of Firms Eliminates Losses:

(i) In Figure 9.4, the representative firm is in a long-run equilibrium. P^e = minimum point of ATC. Normal profits are earned, and hence no economic profits could be generated.

(ii) Suppose demand decreases from D^1 to D^3. Price falls below the minimum point of ATC to p_2 . Economic profits are negative, i.e., firms suffer from economic losses.

(iii) Firms leave the industry and supply falls from S_1, causing price to rise.

(iv) Supply continues to fall while economic loss is being experienced. The reduction of supply will not stop until it reaches the new one S_3 where p^e = the minimum point of ATC again.

(v) In the long run, a smaller quantity is produced and supplied at the same initial price.

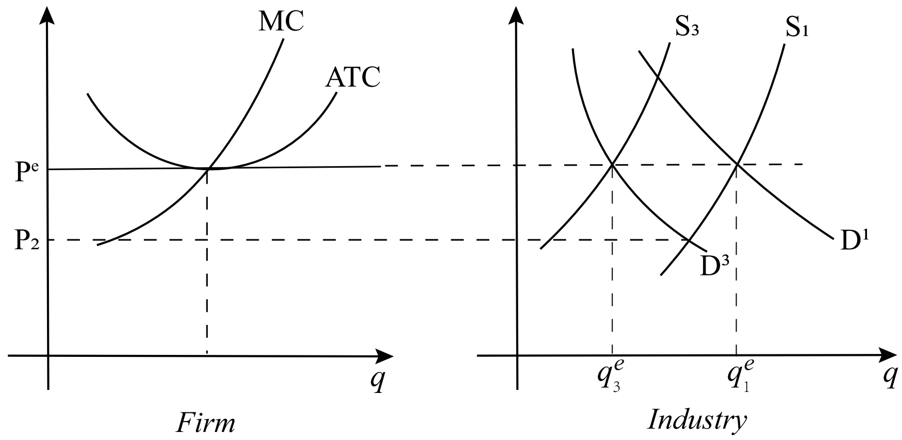

Figure 9.4 Exit of firms eliminates losses

9.3 The Long-Run (LR) Supply Curve

Economists distinguish among three different types of competitive industries, constant-cost, increasing-cost, and decreasing-cost. The distinction depends on how a change in industry output affects the prices of inputs.

Constant-Cost Industry: Horizontal LR Supply Curve

Constant-Cost Industry: Expansion of output does not affect input prices, and consequently the entry and exit of firm do not cause cost curves to shift. The price of output remains constant for all levels of quantity, and hence LR supply is perfectly elastic. The long-run supply curve for a constant-cost industry is, therefore, a horizontal line at a price that is equal to the long-run minimum average cost of production.

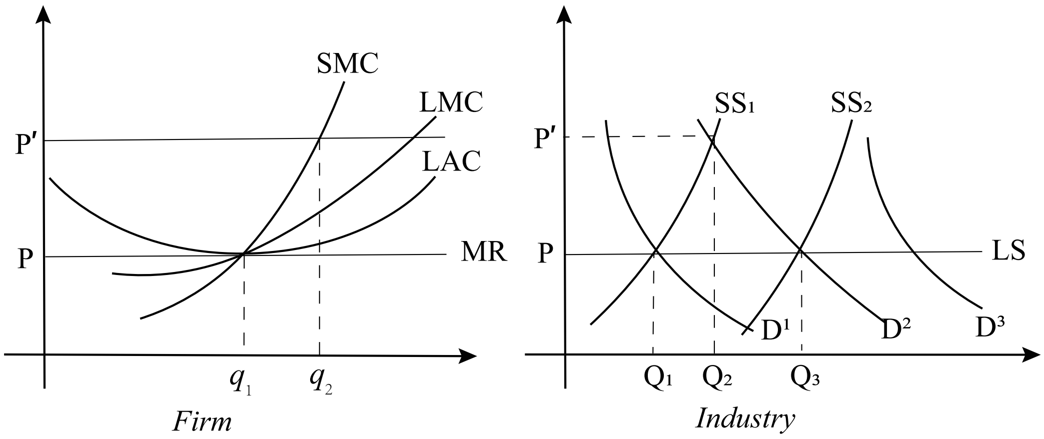

Figure 9.5 The long-run supply curve for a constant-cost industry

Increasing-Cost Industry: Upward-Sloping LR Supply Curve

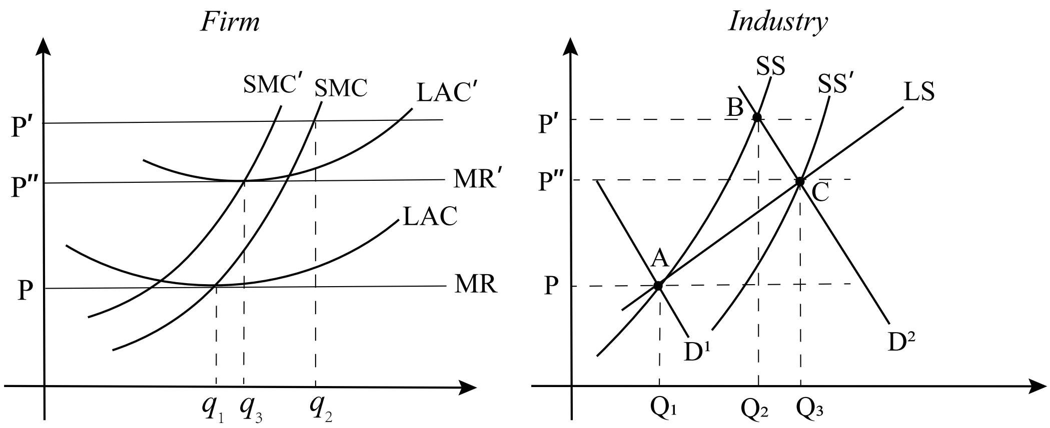

Figure 9.6 The long-run supply curve for an increasing-cost industry

Increasing-Cost Industry: Expansion of output causes input prices to rise as demand for them grow, and consequently the entry/exit of firm makes cost curves shift upward/downward. As firms enter the industry, they compete for scarce resources. Therefore, prices of input resources rise as do production costs of firms. So cost curves shift upward (ATC and MC). Hence industry supply curve is upward-sloping.

Decreasing-Cost Industry: Downward-Sloping LR Supply Curve

Decreasing-Cost Industry: Expansion of output causes input prices to decrease as demand for them grow, and consequently the entry/exit of firm makes cost curves shift downward/upward. Hence, a decreasing-cost industry has a downward-sloping long-run supply curve. This means the expansion of output by the industry in some way lowers the cost curves of individual firms. The entry of firms may result in lower per unit cost.

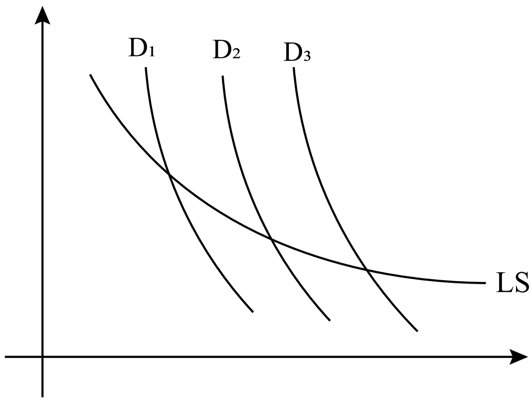

Figure 9.7 The long-run supply curve for a decreasing-cost industry

Example 9.1 We saw that the supply of coffee is extremely elastic in the long run. The reason is that land for growing coffee is widely available and the costs of planting and caring for trees remains constant as the volume grows. Thus, coffee is a constant-cost industry.

The oil industry is an increasing cost industry because there is a limited availability of easily accessible, large-volume oil fields.

Finally, a decreasing-cost industry. In the automobile industry, certain cost advantages arise because inputs can be acquired more cheaply as the volume of production increases.

Long-Run Elasticity of Supply

The long-run elasticity of industry supply is defined in the same way as the short-run elasticity: It is the percentage change in output ($\Delta Q/Q$) that results from a percentage change in price ($\Delta P/P$).

In a constant-cost industry, the long-run supply curve is horizontal, and thus the long-run supply elasticity is infinitely large. (A small increase in price will induce an extremely large increase in output.) In an increasing-cost industry, however, the long-run supply elasticity will be positive but finite.

Because industries can adjust and expand in the long run, we would expect that ***long-run elasticities of supply is generally larger than short-run elasticities***. The magnitude of the elasticity will depend on the extent to which input costs increase as the market expands. For example, an industry that depends on inputs that are widely available will have a more elastic long-run supply than will an industry that uses inputs in short supply.

9.4 Gains and Losses from Government Policies

We now return to supply-demand analysis to show how it can be applied to a wide variety of problems, including situations in which: a consumer faced with a purchasing decision; a firm faced with a long-range planning problem; a government agency that has to design a policy and evaluate its likely impact. We also use consumer and producer surplus to demonstrate the efficiency of a competitive market.

9.4.1 Consumer and Producer Surplus

Consumer surplus measures the total benefit to all consumers, is the shaded area between the demand curve and the market price.

Consumer A would pay \$10 for a good whose market price is \$5 and therefore enjoys a benefit of $5. Consumer B enjoys a benefit of $2, and Consumer C, who values the good at exactly the market price, enjoys no benefit.

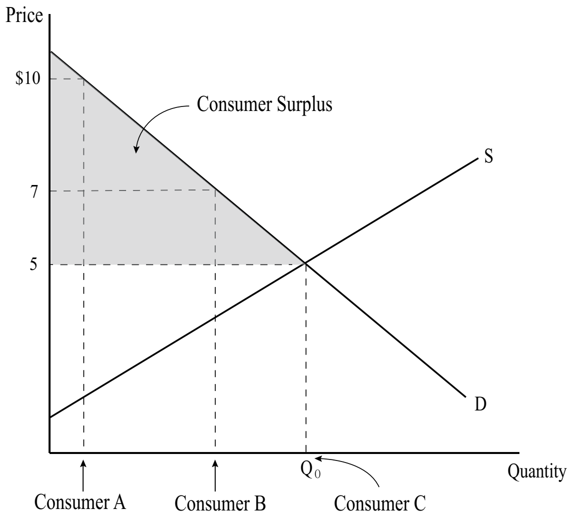

Figure 9.8 Consumer surplus

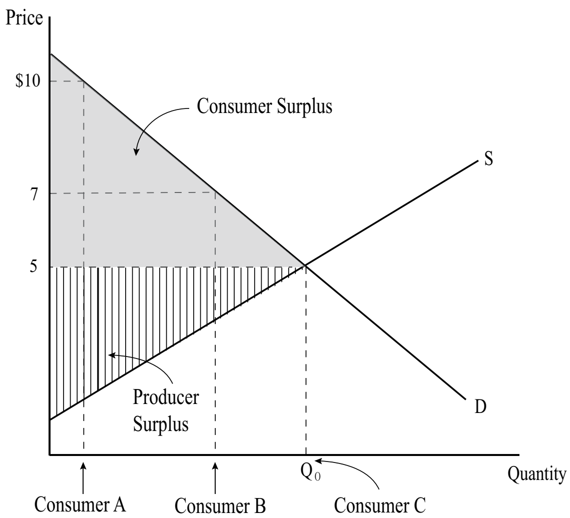

Figure 9.9 Consumer and producer surplus

Producer surplus measures the total profits of producers, plus rents to factor inputs. It is the benefit that lower-cost producers enjoy by selling at the market price, shown by the shaded area between the supply curve and the market price.

Together, consumer and producer surplus measure the **welfare benefit of a competitive market** as shown in Figure 9.9.

Deadweight loss: Net loss of total (consumer plus producer) surplus.

Welfare Effects Gains and Losses to Consumers and Producers

The price of a good has been regulated to be no higher than ceiling price $P_{\max}$, which is below the market-clearing price P_0. The gain to consumers is the difference between rectangle A and triangle B. The loss to producers is the sum of rectangle A and triangle C. Triangles B and C together measure the deadweight loss from price controls.

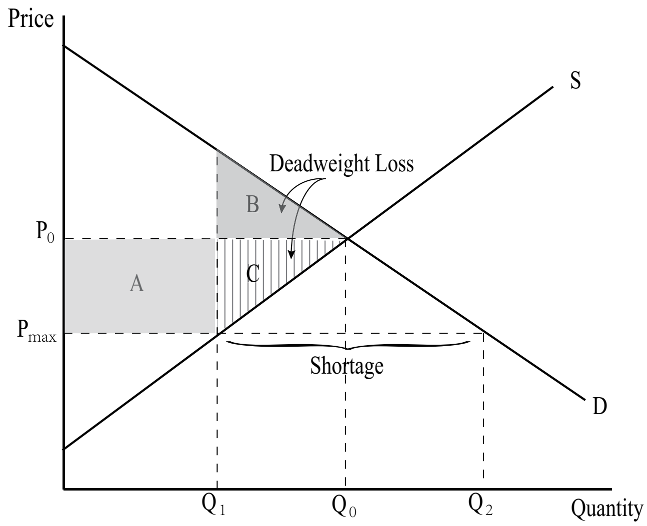

Figure 9.10 Change in consumer and producer surplus from price control

If demand is sufficiently inelastic, triangle B can be larger than rectangle A. In this case, consumers suffer a net loss from price controls as shown in Figure 9.11.

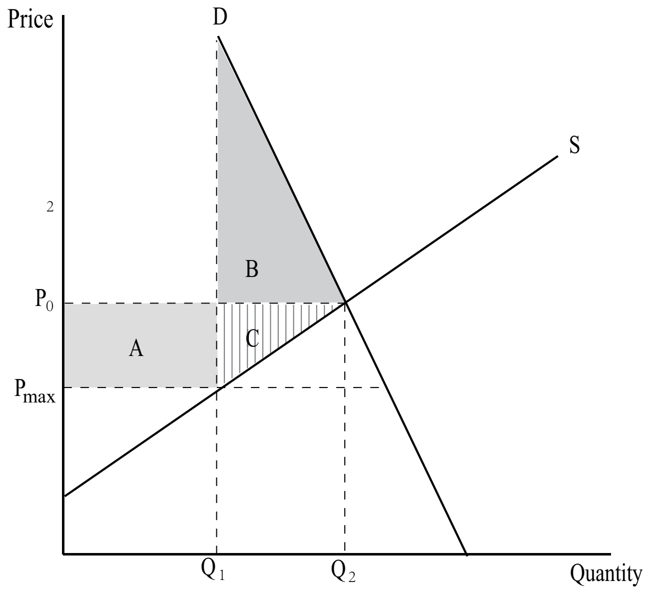

Figure 9.11 Effect of price control when demand is inelastic

9.4.2 The Efficiency of Competitive Markets

Economic efficiency: Maximization of aggregate consumer and producer surplus.

Market failure: Situation in which an unregulated competitive market is inefficient because prices fail to provide proper signals to consumers and producers.

There are two important instances in which market failure can occur:

- Externalities
- Lack of Information

Externality: Action taken by either a producer or a consumer which affects other producers or consumers but is not accounted for by the market price.

Market failure can also occur when consumers lack information about the quality or nature of a product and so cannot make utility-maximizing purchasing decisions. Government intervention (e.g., requiring "truth in labeling") may then be desirable.

Welfare Loss When Price Above Market-Clearing Level

When price is regulated to be no lower than P_2, and then only Q_3 will be demanded. If Q_3 is produced, the deadweight loss is given by triangles B and C. At price P_2, producers would like to produce more than Q_3. If they do, the deadweight loss will be even larger, as shown in Figure 9.12.

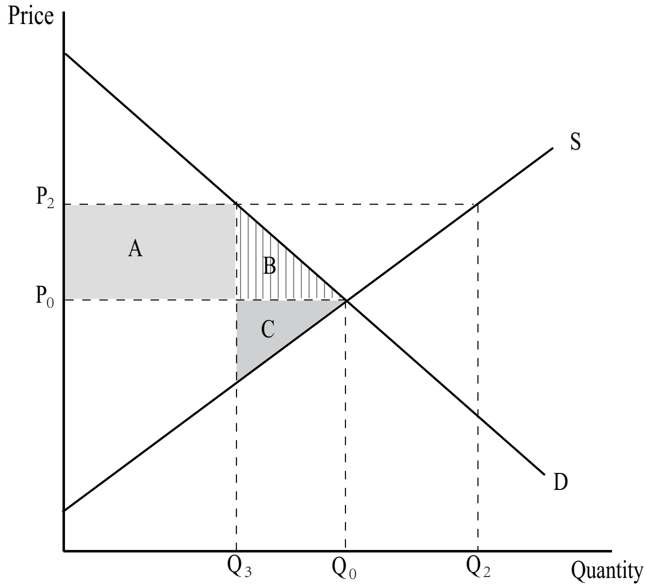

Figure 9.12 Welfare loss when price is held above market-clearing level

9.5 Minimum Prices (Price Floor)

Minimum price is also known the **price floor** or **price support**, which is the price higher than the market price. Suppose that price is regulated to be no lower than $P_{\min}$ as in Figure 9.13. Producers would like to supply Q_2, but consumers will buy only Q_3. Then the amount $Q_2 - Q_3$ will go unsold and the change in producer surplus will be $A - C - D$. In this case, producers as a group may be worse off.

The total change in consumer surplus is:

$$\Delta CS = -A - B.$$

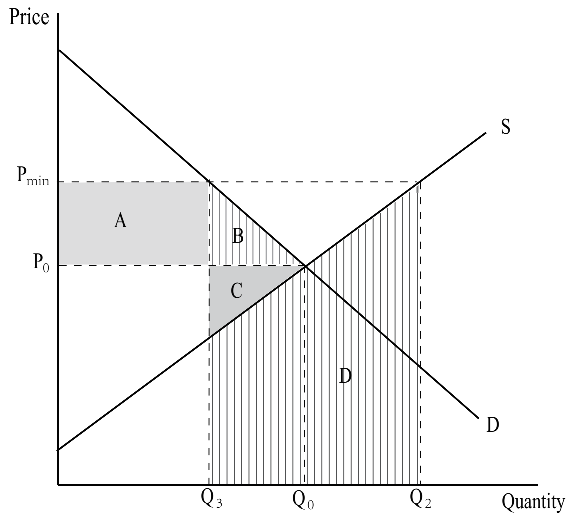

Figure 9.13 Price minimum

The total change in producer surplus is

$$\Delta PS = A - C - D.$$

Minimum Wage

Although the market-clearing wage is w_0, firms are not allowed to pay less than $w_{\min}$.

This results in unemployment of an amount $L_2 - L_1$ and a deadweight loss given by triangles B and C in Figure 9.14.

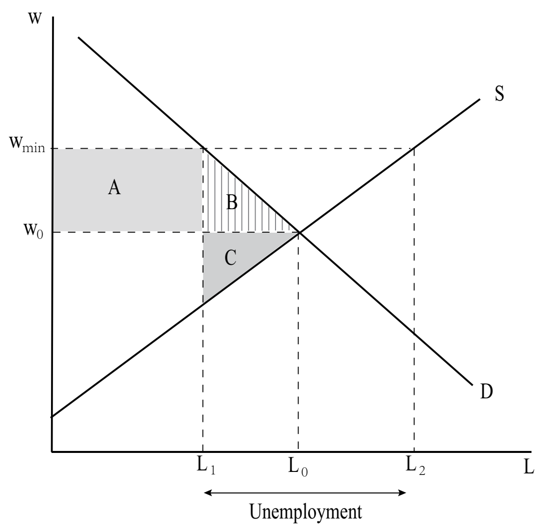

Figure 9.14 Deadweight loss from minimum wage is triangles B and C

Example 9.2 (Airline Regulation) Airline deregulation in 1981 led to major changes in the industry. Some airlines merged or went out of business as new ones entered. Although prices fell considerably (to the benefit of consumers), profits overall did not fall much.

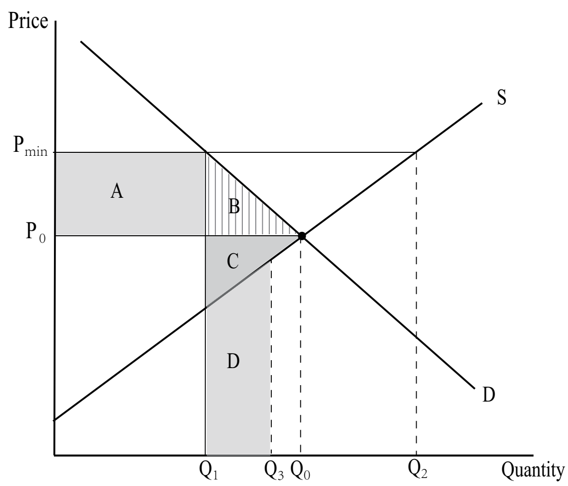

Figure 9.15 Effect of airline regulation by the Civil Aeronautics Board

At price $P_{\min}$, airlines would like to supply Q_2, well above the quantity

Q_1 that consumers will buy.

When they supply Q_3, trapezoid D is the cost of unsold output. Airline profits may have been lower as a result of regulation because triangle C and trapezoid D can together exceed rectangle A. In addition, consumers lose $A + B$.

Part IV

Market Structure and Competitive Strategy

Part 3 examines a broad range of markets and explains how the pricing, investment, and output decisions of firms depend on market structure and the behavior of competitors.

Chapter 10

Monopoly and Monopsony

While perfect competition is characterized by many firms selling in the same market, monopoly is characterized by only one firm selling in a given market. In this chapter, we explain how a monopoly determines price and output, and compare the results with those of the competitive industry. We will then discuss monopsony which is characterized by only one buyer of a good.

10.1 The Nature of Monopoly

Monopoly: A form of market structure in which there is only one seller of some product that has no close substitutes. As a result, the monopoly is the industry because it is the only producer in the market.

Monopoly is the opposite of perfect competition since there is no competition. The monopoly need not be concerned with the possibility that other firms may undercut its price.

Monopsony: Market with only one buyer.

Market power (also known as the **monopoly power** for a seller or **monopsony power** for a buyer): Ability of a seller or buyer to affect the price of a good.

10.2 Sources of Monopoly Power

How does a monopoly power come about?

(1) **Exclusive ownership of a unique resource**. E.g., Debeers Co. of South Africa owns most of the world's diamond mines.

(2) **Economies of scale**. If large economies of scale exist then a firm's LRATC curve will fall over a suitable range as output is increased. The first firm to enter this industry has a competitive advantage because it can take advantage of low per unit costs at higher levels of output, whereas a new firm would have higher per unit costs producing at low levels of output. Thus, the existing firm could charge a price lower than those new firms could afford. Therefore, rivals will not enter the market and monopoly power will be maintained. Such kind of monopoly is known as the **natural monopoly**, e.g., public utilities, telephone services.

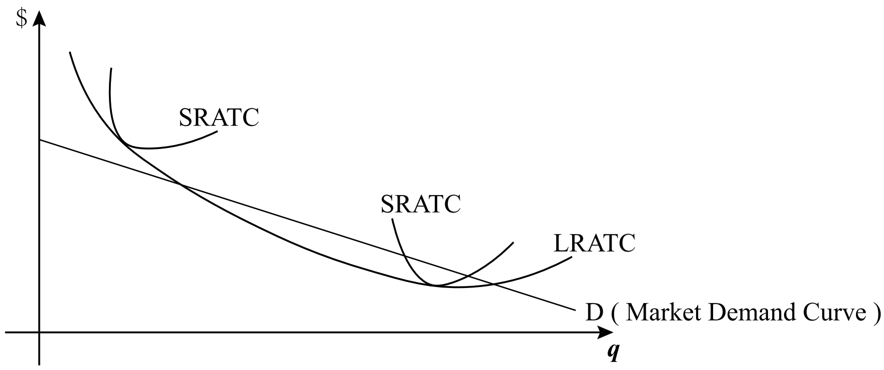

Figure 10.1 LRATC curve under monopoly

(3) **Government-granted monopoly**: The government can grant an exclusive right to produce and sell a product or service. The government can grant such monopoly power through patents, licenses, copyrights, or exclusive franchises.

- ***patents***: give inventor a monopoly position for lifetime of patent–17 years after the patent was granted in the U.S.A. E.g., IBM, Xerox.
- ***copyrights***: give writers and composers exclusive legal controls over production and reproduction of their work for 70 years after the death of the author.
- ***licenses***: limit the number of producers but rarely give monopoly power. E.g., licenses are needed to practice medicine, law, cut hair, or sell liquor.
- ***public utilities***: competition is impractical so industries are given exclusive franchise by the government. If firms share the market none would be able to take advantage of the large economies of scale. If only one firm supplied in the market it could take advantage of the lower per unit costs. In return for the granted monopoly position, the government is allowed to regulate the price of the product.

(4) **Cooperation/Collusion**: Firms may not compete aggressively when they know competition will significantly reduce their profits. They may even collude (in violation of the antitrust laws), agreeing to limit output and raise prices. Other things being equal, monopoly power will be larger when they cooperate. Collusion can generate substantial monopoly power.

10.3 The Monopoly's Demand and Marginal Revenue Curves

Under perfect competition each firm is a price taker and thus faces a horizontal demand curve. By contract, the monopolist comprises the entire industry. Thus, the firm's (monopolist's) demand curve is industry demand curve. Then a monopolist's demand curve is downward sloping—i.e., the firm is a price "maker".

Recall that

- TR: Total Revenue $= P \times Q$;
- MR: Marginal Revenue $= \frac{\Delta TR}{\Delta Q}$;
- AR: Average Revenue $= \frac{P \times Q}{Q} = P$.

Since the demand curve is always the AR curve (i.e., $P = AR$), which is downward sloping for a monopolist, ***marginal revenue must be less than average revenue***, and thus we generally have

$$MR < P.$$

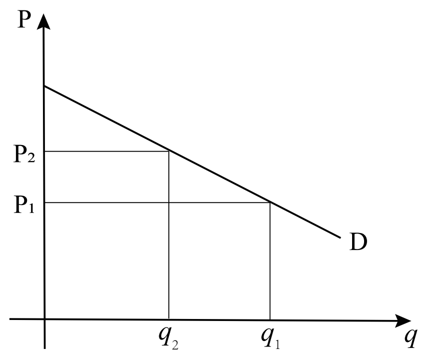

Figure 10.2 Downward-sloping demand curve under monopoly

Indeed, since

$$MR = \frac{\Delta TR}{\Delta Q} = \frac{P\Delta Q + Q\Delta P}{\Delta Q} = P + Q\frac{\Delta P}{\Delta Q}, \tag{10.1}$$

$Q > 0$ and $\frac{\Delta P}{\Delta Q} < 0$, again we have $MR < P$.

Intuitively, the extra revenue from an incremental unit of quantity has two components: (1) producing one extra unit and selling it at P brings in revenue P; and (2) because the firm faces a downward-sloping demand curve, producing and selling this extra unit also results in a small drop in price $\Delta P/\Delta Q$, which reduces the revenue from all units sold (i.e., a change in revenue $q \times \Delta P/\Delta Q$).

Example 10.1 Suppose that the demand curve facing a monopolist is given by

$$P = 11 - Q.$$

Q	P=AR	TR	MR
0	\$11	0	–
1	10	10	10
2	9	18	8
3	8	24	6
4	7	28	4
5	6	30	2
6	5	30	0
7	4	28	-2
8	3	24	-4
9	2	18	-6
10	1	10	-8

For a general linear demand function

$$P = a + bq,$$

$TR = P \times q = aq + bq^2$, and then we have

$$MR = \frac{\Delta TR}{\Delta q} = a + 2bq.$$

Therefore, the slope of the MR curve is $2b$, which is exactly twice the slope of the demand function, b.

10.4 Monopolist's Decision

10.4.1 Monopolist's Profit Maximization

The monopolist can affect both price and quantity. It has both pricing and output policies which are not independent.

Conditions for monopolist's profit maximization:

1) $P \geq AVC$ (necessary condition for producing);

2) $MR = MC$ (how much will it be produced).

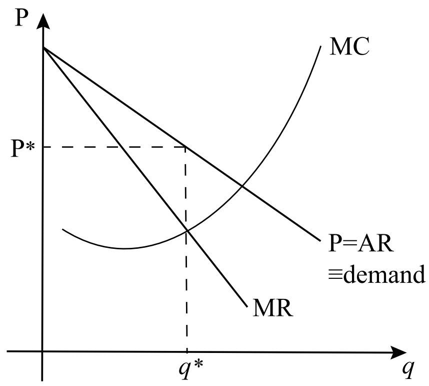

Figure 10.3 Monopolist's profit-maximizing output

We can express profit maximization, loss minimization, shutdown, and normal profit graphically, which depend on relationships among price P, average total cost ATC, and average variable cost AVC.

Profit Maximization: $P > ATC > AVC$

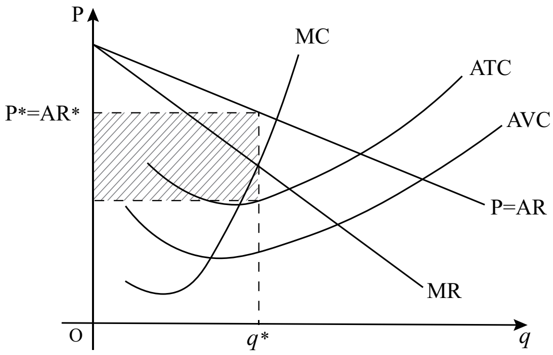

Figure 10.4 Profit maximization for a monopolist

1. Use $MR = MC$ to determine q^*.
2. Use q^* and $P = AR$ (= demand) to decide P^*.
3. $TR = AR \times q^* = P^* \times q^*$.
4. Use q^* and ATC to get TC. $TC = ATC \times q^*$.
5. Profit $= TR - TC$.

Loss Minimization: $ATC > P > AVC$

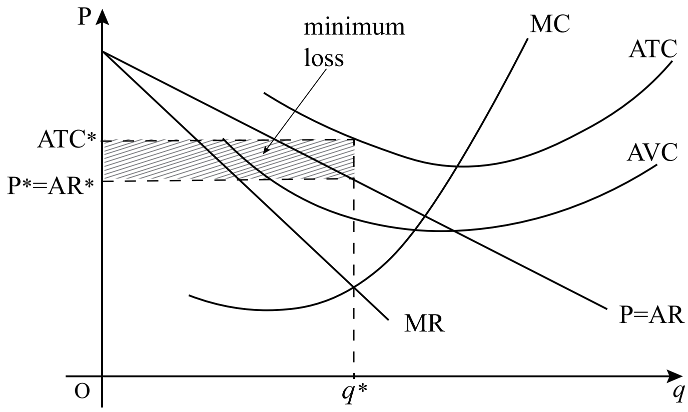

Figure 10.5 Loss minimization for a monopolist

Production Shutdown: $ATC > AVC > P$

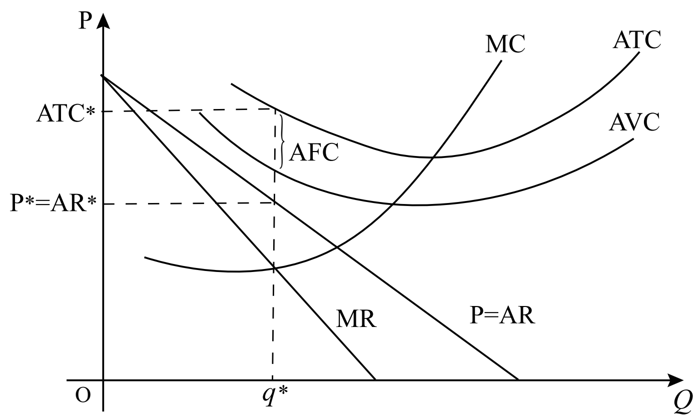

Figure 10.6 Production shutdown for a monopolist

$q^* = 0$ is the loss-minimizing level of output.

Normal Profit: $ATC = P$

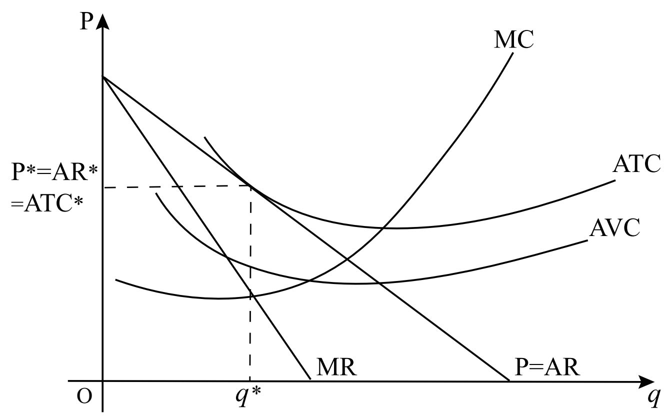

Figure 10.7 Normal profit for a monopolist

$TR = TC \Rightarrow$ Economic Profit $= 0 \Rightarrow$ a normal profit is earned.

Example 10.2 A monopolist can produce at a constant average (and marginal) cost of

$$AC = MC = 2.$$

It faces a market demand curve given by

$$Q = 10 - P.$$

The revenue for the monopolist is

$$TR = PQ = (10 - Q)Q = 10Q - Q_1^2.$$

Then, the marginal revenue is given by

$$MR = \Delta R / \Delta Q = 10 - 2Q.$$

Setting $MR = MC$ and solving for Q, the monopolist's profit-maximizing quantity, profit, and profit are, respectively, given by

$$Q^* = 4, \tag{10.2}$$

$$P^* = 6, \tag{10.3}$$

and

$$\pi = TR - TC = Q^*(P^* - AC) = 16. \tag{10.4}$$

Long-Run: Unlike a competitive firm, a monopolist may earn positive economic profits in the long run since barriers to competition prevent new firms from entering the industry. (Under perfect competition this involved a shift in the supply curve which reduced the equilibrium price and eventually reduced economic profits to zero.)

10.4.2 The Multiplant Firm

Suppose that a firm has two plants. What should its total output be, and how much of that output should each plant produce? We can find the answer intuitively in two steps.

- ***Step 1***. Whatever the total output, it should be divided between the two plants so that marginal cost is the same in each plant. Otherwise, the firm could reduce its costs and increase its profit by reallocating production.

- ***Step 2***. We know that total output must be such that marginal revenue equals marginal cost. Otherwise, the firm could increase its profit by raising or lowering total output.

We can derive this result algebraically. Let Q_1 and C_1 be the output and cost of production for Plant 1, Q_2 and C_2 be the output and cost of production for Plant 2, and $Q_T = Q_1 + Q_2$ be total output. Then profit is

$$\Pi = PQ_T = PQ_1 + PQ_2 - c_1(Q_1) - c_2(Q_2),$$

and thus we have

$$MR = MC_1 = MC_2.$$

Therefore, a firm with two plants maximizes profits by choosing output levels Q_1 and Q_2 so that marginal revenue MR (which depends on total output) equals marginal costs for each plant, MC_1 and MC_2, as depicted in Figure 10.8.

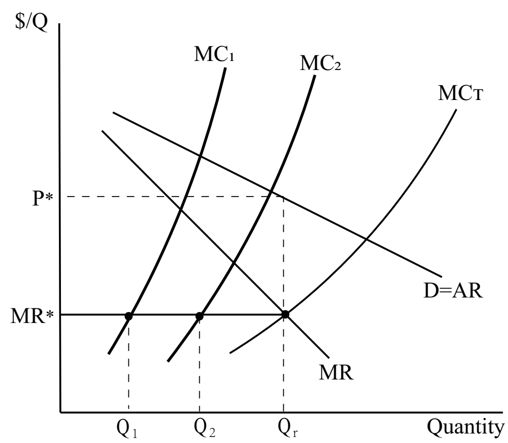

Figure 10.8 Producing with two plants

10.4.3 Relationships between MR, Price, and E_d^P

Since

$$MR = P + Q\frac{\Delta P}{\Delta Q} = P\left(1 + \frac{Q}{P} \times \frac{\Delta P}{\Delta Q}\right) = P\left(1 + \frac{1}{E_d^P}\right),$$

we have

$$MR = P\left(1 + \frac{1}{E_d^P}\right). \tag{10.5}$$

Therefore, ***the less elastic its demand curve, the more monopoly power a firm has***. The ultimate determinant of monopoly power is therefore the firm's elasticity of demand.

By Formula (10.5), we have the following facts:

1. When the elasticity of demand is infinity (a horizontal demand curve), then $MR = P$:

$$MR = P\left(1 - \frac{1}{\infty}\right) = P.$$

2. When demand is unit elastic ($E_d^P = -1$), then $MR = 0$:

$$MR = P\left(1 - \frac{1}{1}\right) = 0.$$

3. When demand is elastic ($|E_d^P| > 1$), then $MR > 0$; e.g., let $E_d^P = -2$:

$$MR = P\left(1 - \frac{1}{2}\right) = \frac{1}{2}P > 0.$$

4. When demand is inelastic ($|E_d^P| < 1$), then $MR < 0$; e.g., let $E_d^P = -\frac{1}{2}$:

$$MR = P\left(1 - \frac{1}{1/2}\right) = P(1-2) = -P < 0.$$

Factors affecting the elasticity of demand of monopolist

- ***The elasticity of market demand***. Because the firm's own demand will be at least as elastic as market demand, the elasticity of market demand limits the potential for monopoly power.

- ***The number of firms in the market***. If there are many firms, it is unlikely that any one firm will be able to affect price significantly.
- ***The interaction among firms***. Even if there are only few firms in the market, each firm will be unable to profitably raise price very much if the rivalry among them is aggressive, with each firm trying to capture as much of the market shares as it can.

It is important to distinct the elasticity of market demand and the elasticity of demand of an individual firm which could be much higher.

Example 10.3 (Elasticity of Demand for Soft Drinks) Soft drinks provide a good example of the difference between a market elasticity of demand and a firm's elasticity of demand.

A recent review of statistical studies found that the market elasticity of demand for soft drinks is between -0.8 and -1.0.6 That means that if all soft drink producers increased the prices of all of their brands by 1 percent, the quantity of soft drinks demanded would fall by 0.8 to 1.06 percent.

The demand for any individual soft drink, however, will be much more elastic, because consumers can readily substitute one drink for another. Although elasticities will differ across different brands, studies have shown that the elasticity of demand for, say, Coca Cola is around -5. In other words, if the price of Coke were increased by 1 percent but the prices of all other soft drinks remained unchanged, the quantity of Coke demanded would fall by about 5 percent.

Students-and business people-sometimes confuse the market elasticity of demand with the firm (or brand) elasticity of demand. Make sure you understand the difference.

10.4.4 A Rule of Thumb for Pricing

With limited knowledge of average and marginal revenue, we can derive a rule of thump that can be more easily applied in practice. From (10.5), we know the marginal revenue can be expressed as

$$MR = P\left(1 + \frac{1}{1/E_d^P}\right).$$

Since $MR = MC$ at the profit-maximizing output, we have

$$P\left(1 + \frac{1}{1/E_d^P}\right) = MC, \tag{10.6}$$

which can be rearranged to get the so-called **markup equation**:

$$P = \frac{MC}{1 + 1/E_d^P}. \tag{10.7}$$

Since $E_d^P < 0$, in order to ensure that the price is non-negative, the firm must produce ***within the arrangement of the elastic demand***. Therefore, we have $[1 + 1/E_d^P] \leqq 1$ or equivalently, $\frac{1}{1+1/E_d^P} \geqq 1$, which implies that the price is not less than the marginal cost at profit maximization.

This formula is very useful, and provides a basic principle of pricing for any market structure. It reveals that the price equals the marginal cost multiplied by the markup factor $\frac{1}{1+1/E_d^P}$, which is a decreasing function of the price elasticity of demand E_d^P. That is, given a marginal cost, the optimal pricing of the product is inversely proportional to the price elasticity of demand, which means that the smaller is the elasticity, the bigger is the market power and thus the higher is the price. When the market is perfectly competitive, E_d^P is infinitely large, we particularly have $P = MC$.

Example 10.4 (Markup Pricing) Although the elasticity of market demand for food is small (about -1), no single supermarket can raise its prices very much without losing customers to other stores.

The elasticity of demand for any one supermarket is often as large as

-10. We then find that by markup equation (10.7),

$$P = MC/(1-0.1) = MC/(0.9) = (1.11)MC,$$

which is about 11 percent above marginal cost.

Small convenience stores typically charge higher prices because its customers are generally less price sensitive. Because the elasticity of demand for a convenience store is about -5, the markup equation (10.7) implies that its prices should be about 25 percent above marginal cost. With designer jeans, demand elasticities in the range of -2 to -3 are typical. This means that price should be 50 to 100 percent higher than marginal cost.

Example 10.5 (Astra-Merck Prices Prilosec) In 1995, Prilosec, represented a new generation of antiulcer medication. Prilosec was based on a very different biochemical mechanism and was much more effective than earlier drugs. By 1996, it had become the best-selling drug in the world and faced no major competitor.

Astra-Merck was pricing Prilosec at about \$3.50 per daily dose. The marginal cost of producing Prilosec is only about 30 to 40 cents per dose.

The price elasticity of demand should be in the range of roughly -1.0 to -1.2. Setting the price at a markup exceeding 400 percent over marginal cost is consistent with our rule of thumb for pricing.

10.4.5 Measuring Monopoly Power

The important distinction between a perfectly competitive firm and a firm with monopoly power: For the competitive firm, price equals marginal cost; for the firm with monopoly power, price exceeds marginal cost.

Lerner Index of Monopoly Power: Measure of monopoly power calculated as excess of price over marginal cost as a fraction of price:

$$L = \frac{P - MC}{P},$$

which is also equal to minus the inverse of E_d^p by equation (10.6):

$$L = \frac{P - MC}{P} = -\frac{1}{E_d^p}.$$

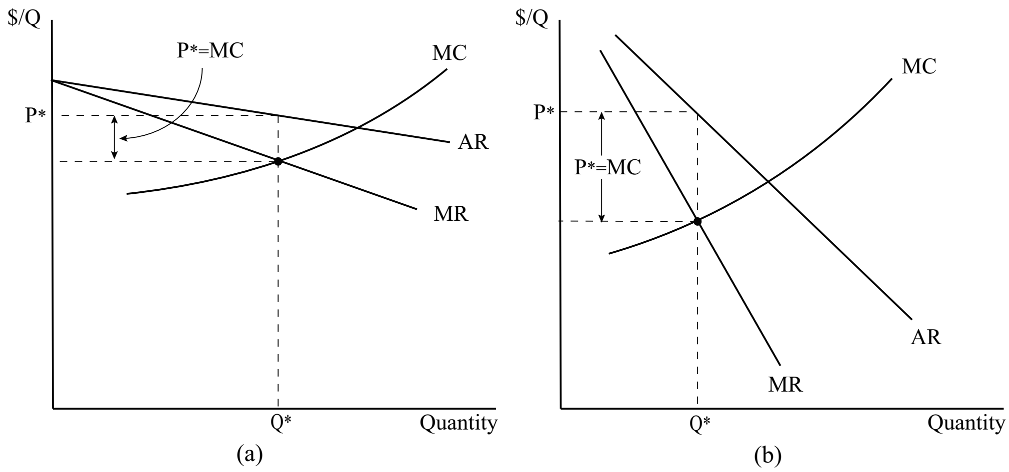

Figure 10.9 Elasticity of demand and price markup

Thus, if the firm's demand is elastic, as in (a) of Figure 10.9, the markup is small and the firm has little monopoly power. The opposite is true if demand is relatively inelastic, as in (b) of Figure 10.9.

10.5 Further Implications of Monopoly Analysis

Monopoly Has No Supply Curve

Under perfect competition we noticed that at any given price there is a unique quantity of output that the firm is willing to supply. Conversely, at any given output level there is a unique price that makes the firm be willing to supply that output level. This relationship does **not** exist for the case of a monopolist, and hence monopoly has no supply curve. The reason is that ***a monopolist's output decision depends not only on marginal cost but also on the demand curve***.

Consider two possible demand curves below in Figure 10.10: MR_1 and MR_2 intersect MC at the same point:

- If demand curve is D_1, then price is P_1.
- If demand curve is D_2, then price is P_2.

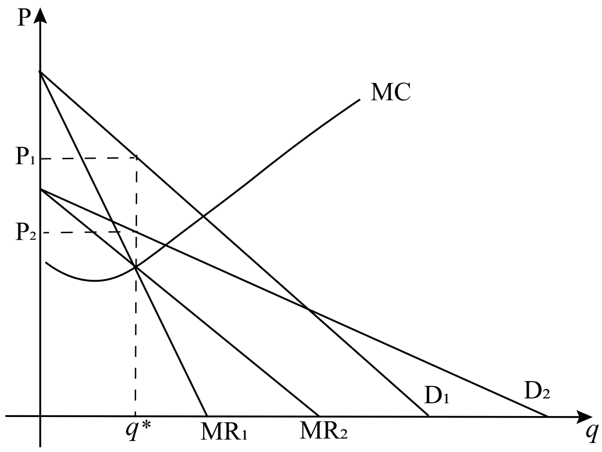

Figure 10.10 No supply curve: A change of demand only changes the price charged

Thus, a MC curve cannot be supply curve for a monopolist. Therefore, a ***monopoly has no supply curve***. In other words, there is no one-to-one relationship between price and the quantity produced. As a result, shifts in demand do not trace out the series of prices and quantities that correspond to a competitive supply curve. Instead, shifts in demand can lead to changes in price with no change in output, changes in output with no change in price, or changes in both price and output.

However, the absence of a monopoly supply curve does not mean that we are unable to analyze the output choice of a monopoly.

10.6 Monopoly versus Perfect Competition

Perfectly Competitive Industry: $S = \sum_i MC_i$, $S = D \Rightarrow P^c, q^c$.

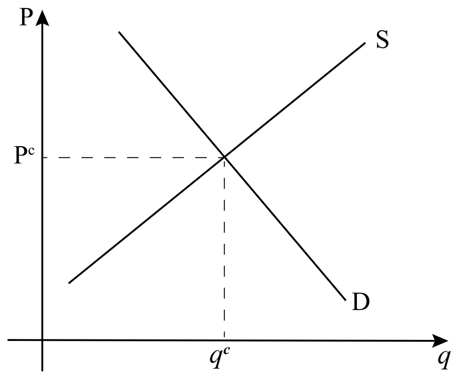

Figure 10.11 Market equilibrium under perfect competition

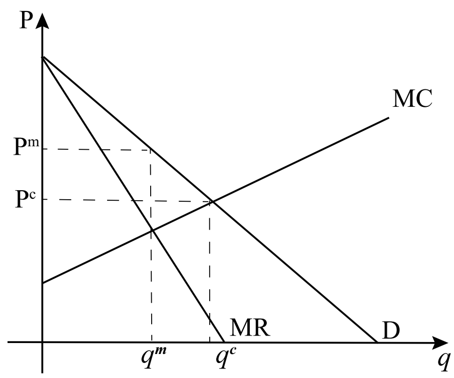

Figure 10.12 Market equilibrium under monopoly

Monopoly (a single firm takes over all firms): $MR = MC \Rightarrow q^m, P^m$.

Therefore, when an industry is a monopoly consumers pay a higher price and receive less than would be the case under perfect competition.

10.7 The Social Cost of Monopoly Power

10.7.1 Income Distribution Problem

If a competitive industry becomes a monopoly, there will be a change in the distribution of real income among members of society. The monopolist will gain. Consumers will lose because of a higher price they pay.

The higher price reduces the real purchasing power, or real income, of the consumers. Thus, by changing a price above average cost, the monopolist gains at the expense of the consumers–a redistribution of income from consumers to the owners of the monopoly.

10.7.2 Inefficient Allocations

Monopoly has another effect that involves a net loss in welfare because it leads to an inefficient allocation. Economists refer to this net loss as a social cost of monopoly.

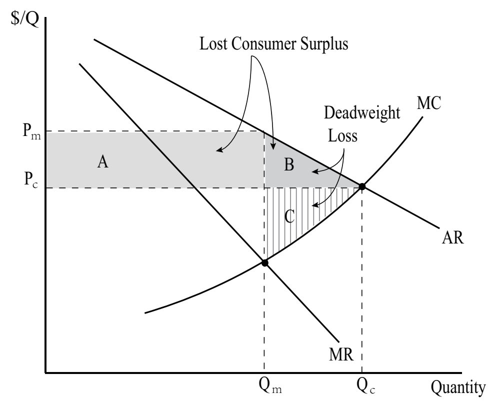

Figure 10.13 Deadweight Loss form Monopoly Power

The shaded rectangle and triangles in Figure 10.13 show changes in consumer and producer surplus when moving from competitive price and quantity, P_c and Q_c, to a monopolist's price and quantity, P_m and Q_m.

Because of the higher price, consumers lose $A + B$ and producer gains $A - C$. The deadweight loss is $B + C$.

Price Regulation

How can we solve the inefficiency issue of monopoly? One way is to regular the price, especially to the competitive price.

If left alone, a monopolist produces Q_m and charges P_m as shown in 10.14. When the government imposes a price ceiling of P_1 the firm's average and marginal revenue are constant and equal to P_1 for output levels up to Q_1. For larger output levels, the original average and marginal revenue curves apply. The new marginal revenue curve is, therefore, the dark purple line, which intersects the marginal cost curve at Q_1.

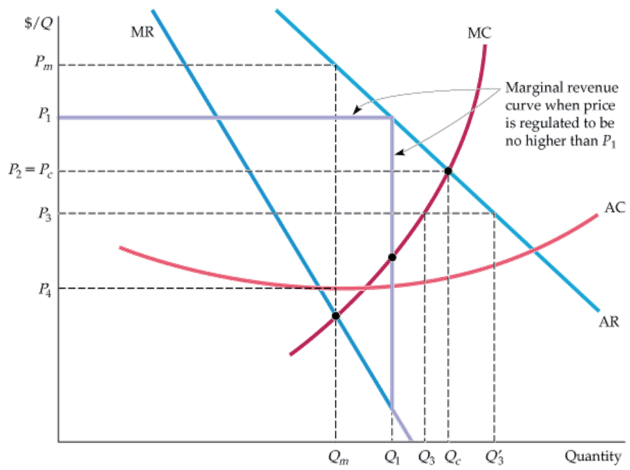

Figure 10.14 Price Regulation

When price is lowered to P_c, at the point where marginal cost intersects average revenue, output increases to its maximum Q_c. This is the output that would be produced by a competitive industry.

Lowering price further, to P_3, reduces output to Q_3 and causes a shortage, $Q'_3 - Q_3$.

10.8 Benefits of Monopoly: Corporate Innovation

Although social costs of monopoly, it is not without certain merit. The most merit of monopoly is that it provides strong incentives for the firm to innovate continuously to conduct research and develop new products.

Corporate innovation leads to monopoly profits, and considerable profits will attract other firms to enter and compete again. In this way, market competition leads to the decrease of profits and firms obtain monopoly profits through innovation, which forms a **repeated cycle of competition-innovation-monopoly-competition**. In this cycle, market competition produces market equilibrium, while innovation disrupts this equilibrium. This repeated game in the market motivates firms to constantly pursue innovation. Through this repeated game process, the market economy maintains long-term vitality, increases social welfare and promotes economic development, revealing the unique beauty and power of the market system.

Of course, in order to encourage innovation, the government should enact intellectual property protection laws. At the same time, to encourage competition and form externalities of technology innovation, the protection given by anti-monopoly laws and intellectual property rights legislation should not be permanent, but rather for a limited time, in case fixed or permanent oligopolies and monopolies may appear.

Thus, competition and monopoly are two sides of the same entity, like supply and demand, the two of which can become an awe-inspiring dialectical unity of opposites under market forces, showing the beauty and power of the market system.

10.9 Monopsony

Monopsony power: Buyer's ability to affect the price of a good.

Oligopsony: Market with only a few buyers.

Marginal value (MV): Additional benefit derived from purchasing one more unit of a good.

Marginal expenditure (ME): Additional cost of buying one more unit of a good.

Average expenditure (AE): Price paid per unit of a good.

10.9.1 Competitive Buyer Compared to Competitive Seller

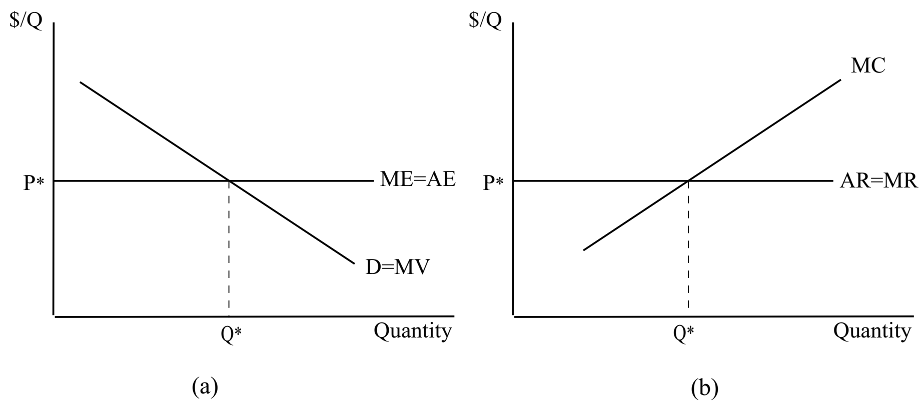

Figure 10.15 Competitive buyer compared to competitive seller

In Figure 10.15.(a), the competitive buyer takes price P* as given, hence ***marginal expenditure and average expenditure are constant and equal***:

$$ME = AV.$$

Quantity purchased is found by price equating marginal value (demand).

In Figure 10.15.(b), the competitive seller also takes price as given. Marginal revenue and average revenue are constant and equal; quantity sold is found by equating price to marginal cost.

10.9.2 Monopsony Buyer

The market supply curve is monopsonist's average expenditure curve AE. Because average expenditure is rising, marginal expenditure lies above it.

In Figure 10.16, the monopsonist purchases quantity Q_m^*, where marginal expenditure and marginal value (i.e., demand) intersect.

The price paid per unit P_m^* is then found from the average expenditure (i.e., supply) curve. In a competitive market, price and quantity, P_c and Q_c, are both higher. They are found at the point where average expenditure (supply) and marginal value (demand) intersect.

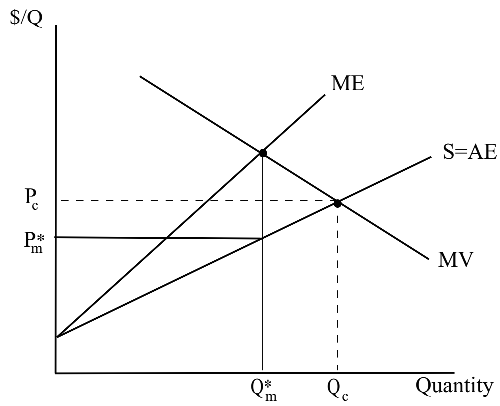

Figure 10.16 Monopsonist buyer

Monopoly and Monopsony

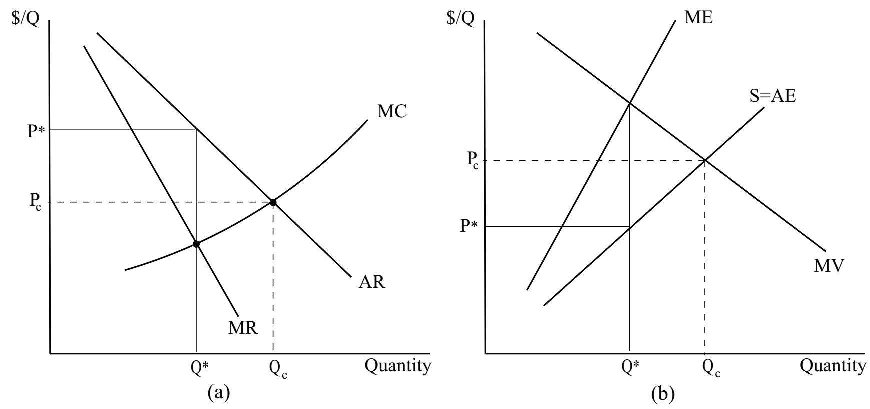

Figure 10.17 Monopoly and Monopsony

These diagrams show the close analogy between monopoly and monopsony.

(a) The monopolist produces where marginal revenue intersects marginal cost. Average revenue (i.e. demand) exceeds marginal revenue, so that price exceeds marginal cost.

(b) The monopsonist purchases up to the point where marginal expenditure intersects marginal value.

Marginal expenditure exceeds average expenditure (i.e., supply), so that marginal value exceeds price.

10.10 Monopsony Power

10.10.1 Monopsony Power: Elastic versus Inelastic Supply

Monopsony power depends on the elasticity of supply. When supply is elastic, as in Figure 10.18.(a), marginal expenditure and average expenditure do not differ by much, so price is close to what it would be in a competitive market.

The opposite is true when supply is inelastic, as in Figure 10.18.(b).

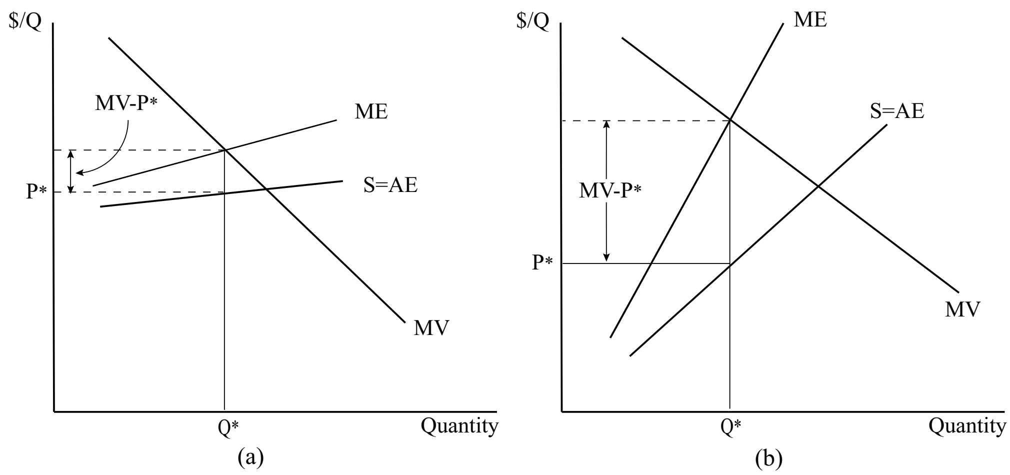

Figure 10.18 Monopsony power: elastic versus inelastic supply

10.10.2 Sources of Monopsony Power

- **Elasticity of market supply**: If only one buyer is in the market-a pure monopsonist, its monopsony power is completely determined by the elasticity of market supply. If supply is highly elastic, monopsony power is small and there is little gain in being the only buyer.

- **Number of buyers**: When the number of buyers is very large, no single buyer can have much influence over price. Thus each buyer faces an extremely elastic supply curve, so that the market is almost completely competitive.

- **Interaction among buyers**: If buyers in a market compete aggressively, they will bid up the price close to their marginal value of the product, and will thus have little monopsony power. On the other hand, if those buyers compete less aggressively, or even collude, prices will not be bid up very much, and the degree of monopsony power might be nearly as high as if there were only one buyer.

10.10.3 The Social Costs of Monopsony Power

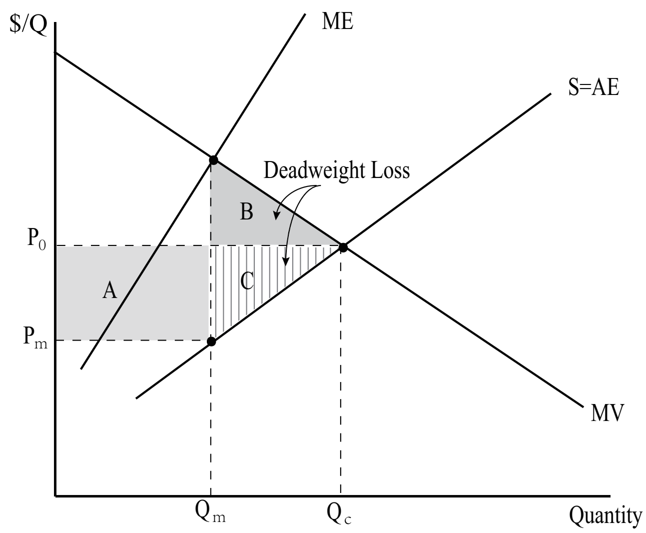

Figure 10.19 Monopsony power: elastic versus inelastic supply

The shaded rectangle and triangles in Figure 10.19 show changes in buyer and seller surplus when moving from competitive price and quantity, P_c

and Q_c, to the monopsonist's price and quantity, P_m and Q_m.

Because both price and quantity are lower, there is an increase in buyer (consumer) surplus given by $A - B$. Producer surplus falls by $A + C$, so there is a deadweight loss given by triangles B and C.

10.10.4 Bilateral Monopsony

Bilateral monopoly: Market with only one seller and one buyer.

It is difficult to predict the price and quantity in a bilateral monopoly. Both the buyer and the seller are in a bargaining situation.

Bilateral monopoly is rare. Although bargaining may still be involved, we can apply a rough principle here: Monopsony power and monopoly power will tend to counteract each other. In other words, the monopsony power of buyers will reduce the effective monopoly power of sellers, and vice versa.

This tendency does not mean that the market will end up looking perfectly competitive, but in general, monopsony power will push price closer to marginal cost, and monopoly power will push price closer to marginal value.

10.11 Limiting Market Power: The Antitrust Laws

Excessive market power harms potential purchasers and raises problems of equity and fairness. In addition, market power reduces output, which leads to a deadweight loss.

In theory, a firm's excess profits could be taxed away, but redistribution of the firm's profits is often impractical.

To limit the market power of a natural monopoly, such as an electric utility company, direct price regulation is the answer.

Antitrust laws: Rules and regulations prohibiting actions that restrain, or are likely to restrain, competition. That is, ***it is designed to promote to***

competition.

It is important to stress that, while there are limitations (such as colluding with other firms), in general, it is not illegal to be a monopolist or to have market power. On the contrary, we have seen that patent and copyright laws protect the monopoly positions of firms that developed unique innovations.

Restricting What Firms Can Do

Parallel conduct: Form of implicit collusion in which one firm consistently follows actions of another.

Predatory pricing: Practice of pricing to drive current competitors out of business and to discourage new entrants in a market so that a firm can enjoy higher future profits.

Enforcement of the Antitrust Laws

The antitrust laws are enforced in three ways:

- Through the Antitrust Division of the Department of Justice.
- Through the administrative procedures of the Federal Trade Commission.
- Through private proceedings.

Example 10.6 (Microsoft) Over the past three decades Microsoft has dominated the software market for personal computers, having maintained over a 90-percent market share for operating systems and for office suites.

In 1998, the Antitrust Division of the U.S. DOJ filed suit, claiming that Microsoft had illegally bundled its Internet browser with its operating system for the purpose of maintaining its dominant operating system monopoly. The court found that Microsoft did have monopoly power in the market for PC operating systems, which it had maintained illegally in violation of Section 2 of the Sherman Act.

In 2004, the European Commission claimed that by bundling the Windows Media Player with the operating system, Microsoft would monopolize the market for media players. Microsoft agreed to offer customers a choice of browsers when first booting up their new operating system, and the case came to a close in 2012.

By 2016, the locus of competition had moved to the smartphone industry, where Microsoft faces strong competition from Google's (Android) and Apple's (iOS) operating systems.

Chapter 11

Pricing with Market Power

11.1 Capturing Consumer Surplus

The pricing of monopoly discussed in previous chapter is uniform price, but in reality, firms commonly adopt differential pricing. The premise of differential pricing is that the purchase of the demander will generate consumer surplus. A monopolist may obtain some, or even all, of the consumer surplus through price discrimination.

Price discrimination: Selling the same good or service at different prices to different buyers.

Monopolists use price discrimination to realize greater profits. They have ability to do this since they are the only supplier of a good. E.g.,

- movie theaters charge different prices for children and adults.
- utility companies charge different rates for businesses and residences.
- senior citizens pay a lower fare for bus ride.

Conditions which aid price discrimination:

- resale not possible;
- must be able to segment the market by classifying buyers in separate, identifiable groups (say, by showing ID);

- monopoly control;
- different demand elasticities: E.g., Mr. A's demand is less elastic, and Mr. B's demand is more elastic. Monopolists can increase TR by increasing price for Mr. A while decreasing price for Mr. B.

There are many broadly adopted forms of price discrimination in practice.

11.2 First-Degree Price Discrimination

Reservation price: Maximum price that a customer is willing to pay for a good.

First degree price discrimination: Practice of charging each customer her reservation price.

11.2.1 Capturing Consumer Surplus by Perfect Price Discrimination

Assume below that $MC = ATC$ which is constant.

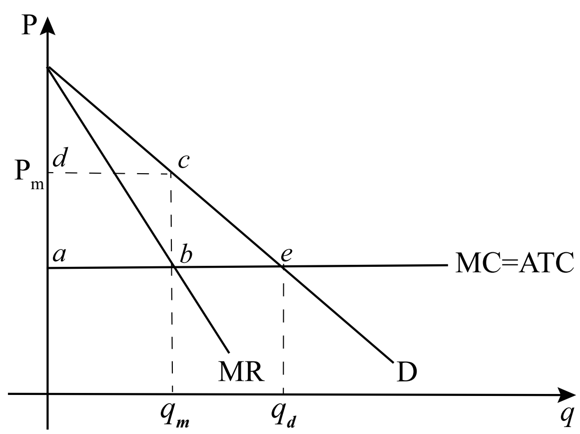

Figure 11.1 Market equilibrium under perfect price discrimination

If the same price is charged for all goods, then profit is maximized at q_m, sold at p_m. Profit = area $abcd$.

If the monopolist uses perfect price discrimination to charge the maximum price buyers are willing to pay for each additional unit, then the profit expands to the area between the demand curve and the marginal cost curve.

11.2.2 Capturing Consumer Surplus by Imperfect Price Discrimination

Firms usually do not know the reservation price of every consumer, but sometimes reservation prices can be roughly identified. In Figure 11.2, six different prices are charged. The firm earns higher profits, but some consumers may also benefit. With a single price P_4^*, there are fewer consumers. The consumers who now pay P_5 or P_6 enjoy a surplus.

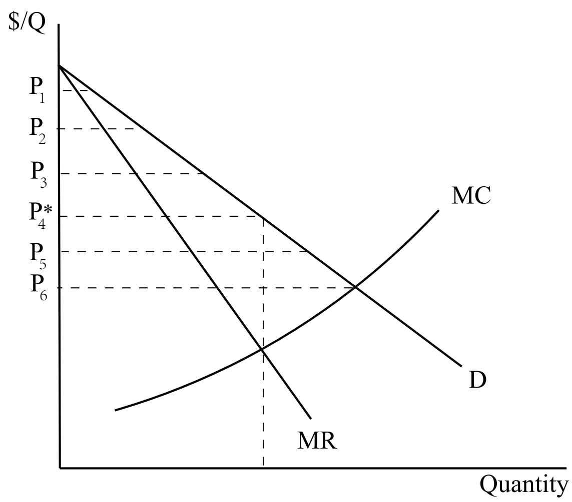

Figure 11.2 First-degree price discrimination in practice

11.3 Second-Degree Price Discrimination

Second-degree price discrimination: Practice of charging different prices per unit for different quantities of the same good or service.

Block pricing: Practice of charging different prices for different quantities or "blocks" of a good.

Different prices are charged for different quantities, or "blocks," of the same good. Here, there are three blocks, with corresponding prices P_1, P_2, and P_3 in Figure 11.3.

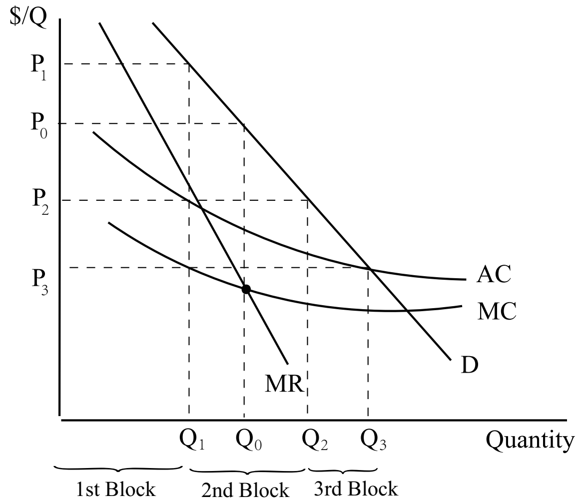

Figure 11.3 Second-degree price discrimination

There are also economies of scale, and average and marginal costs are declining. Second-degree price discrimination can then make consumers better off by expanding output and lowering cost.

11.4 Third-Degree Price Discrimination

Third-degree price discrimination: Practice of dividing consumers into two or more groups with ***separate demand curves*** and charging different prices to each group.

Creating Consumer Groups

If third-degree price discrimination is feasible, how should the firm decide what price to charge each group of consumers?

We know that however much is produced, total output should be divided between the groups of customers so that marginal revenues for each group are equal. We know that total output must be such that the marginal revenue for each group of consumers is equal to the marginal cost of production.

Let P_1 be the price charged to the first group of consumers, P_2 the price charged to the second group, and $C(Q_T)$ the total cost of producing output $Q_T = Q_1 + Q_2$. Total profit is then

$$\Pi = P_1Q_1 + P_2Q_2 - C(Q_T)$$

and thus

$$MR_1 = MR_2 = MC.$$

Therefore, from

$$MR = P(1 + 1/E_d^p),$$

the relative prices are given by

$$\frac{P_1}{P_2} = \frac{(1 + 1/E_2)}{(1 + 1/E_1)}.$$

Consider Figure 11.4. Consumers are divided into two groups, with separate demand curves for each group. The optimal prices and quantities are such that the marginal revenue from each group is the same and equal to marginal cost.

Here group 1, with demand curve D_1, is charged P_1, and group 2, with the more elastic demand curve D_2, is charged the lower price P_2. Marginal cost depends on the total quantity produced Q_T. Note that Q_1 and Q_2 are chosen so that $MR_1 = MR_2 = MC$.

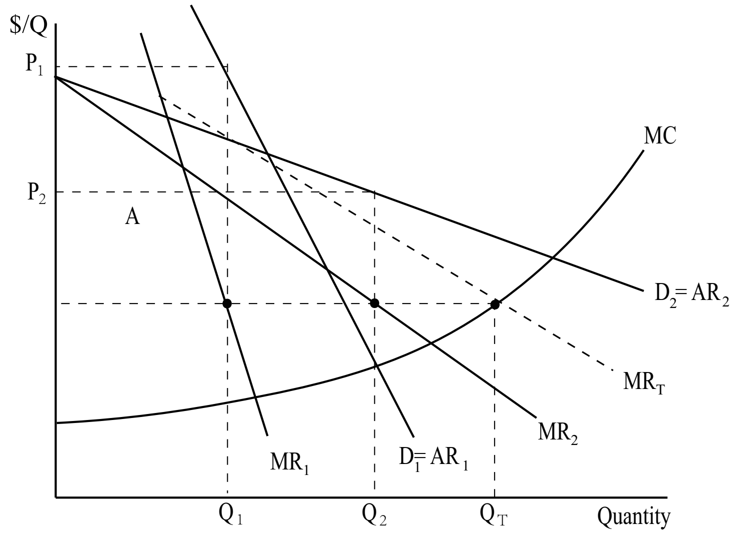

Figure 11.4 Third-degree price discrimination

Example 11.1 (The Economics of Coupons and Rebates) Coupons provide a means of price discrimination.

	Price Elasticity	Price Elasticity
Product	Nonusers	Users
Toilet	-0.60	-0.66
Stuffing/dressing	-0.71	-0.96
Shampoo	-0.84	-1.04
Cooking/salad oil	-1.22	-1.32
Dry mix dinners	-0.88	-1.09
Cake mix	-0.21	-0.43
Cat food	-0.49	-1.13
Frozen entrees	-0.60	-0.95
Gelatin	-0.97	-1.25
Spaghetti sauce	-1.65	-1.81
Creme rinse/conditioner	-0.82	-1.12
Soups	-1.05	-1.22
Hot dogs	-0.59	-0.77

Table 11.1: Price elasticity of demand for users versus nonusers of coupons

Example 11.2 (Airline Fares) Travelers are often amazed at the variety of fares available for round-trip flights.

	Fare Category	Fare Category	Fare Category
Elasticity	First Class	Unrestricted Coach	Discouted
Price	-0.3	-0.4	-0.9
Income	1.2	1.2	1.8

Table 11.2: Elasticities of demand for air travel

Recently, for example, the first-class fare for round-trip flights from New York to Los Angeles was above \$2000; the regular (unrestricted) economy fare was about \$1000, and special discount fares (often requiring the purchase of a ticket two weeks in advance and/or a Saturday night stay over) could be bought for as little as \$200. These fares provide a profitable form of price discrimination. The gains from discriminating are large because different types of customers, with very different elasticities of demand, purchase these different types of tickets. Airline price discrimination has become increasingly sophisticated. A wide variety of fares is available.

11.5 The Two-Part Tariff

Two-part tariff: Form of pricing in which consumers are charged both an entry fee and a usage fee.

It is also commonly used in practice.

11.5.1 Two-Part Tariff with a Single Consumer

The consumers have market demand curve D as shown in Figure 11.5. The firm maximizes profit by setting usage fee P equal to marginal cost and entry fee T^* equal to the entire surplus of the consumer.

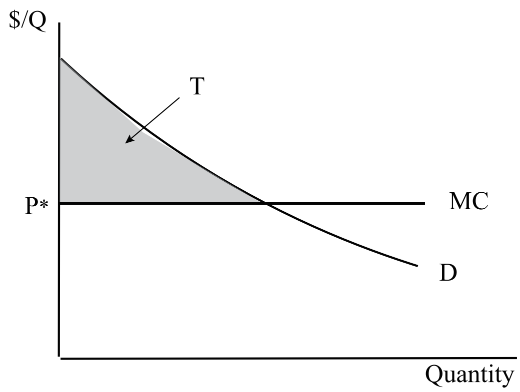

Figure 11.5 Two-Part Tariff with a Single Consumer

11.5.2 Two-Part Tariff with Two Consumer

The profit-maximizing usage fee P^* will exceed marginal cost. The entry fee T^* is equal to the surplus of the consumer with the smaller demand.

The resulting profit is $2T^* + (P^* - MC)(Q_1 + Q_2)$. Note that this profit is larger than twice the area of triangle ABC as shown in Figure 11.6.

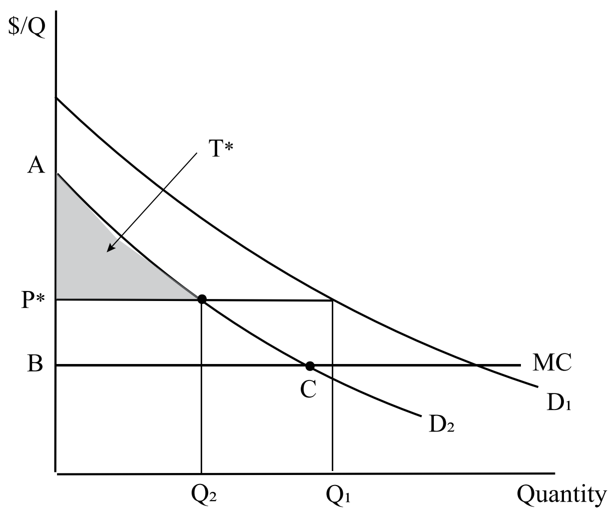

Figure 11.6 Two-Part Tariff with a Single Consumer

Example 11.3 (Pricing Cellular Phone Service) Cellular phone service has traditionally been priced using a two-part tariff: a monthly access fee, which includes some amount of free "anytime" minutes, plus a per-minute charge for additional minutes.

Offering different plans allowed companies to combine third-degree price discrimination with the two-part tariff.

Today, most consumers use their phone not just to make or receive calls but also to surf the Web, read email, and so on. They separate themselves into groups based on their expected data usage, with each group choosing a different plan. Cellular providers have learned that the most profitable way to price their service is to combine price discrimination with a two-part tariff.

Data Usage	Monthly Price	Monthly Access Charge	Overage Fee
A. Verizon	A. Verizon	A. Verizon	A. Verizon
1 GB	$30	$ 20	$ 15/GB
3GB	$45	$ 20	$ 15/GB
6GB	$60	$ 20	$ 15/GB
12GB	$80	$ 20	$ 15/GB
18GB	$100	$ 20	$ 15/GB
B. Sprint	B. Sprint	B. Sprint	B. Sprint
1GB	$20	$ 45	$ none
3GB	$30	$ 45	$ none
6GB	$45	$ 45	$ none
12GB	$60	$ 45	$ none
24GB	$80	$ 45	$ none
C. AT&T	C. AT&T	C. AT&T	C. AT&T
2GB	$30	$ 25	$ 15/GB
5GB	$50	$ 25	$ 15/GB
15GB	$100	$ 15	$ 15/GB
20GB	$140	$ 15	$ 15/GB
25GB	$175	$ 15	$ 15/GB
30GB	$225	$ 15	$ 15/GB

Table 11.3: Cellular data plans (2016)

11.6 Intertemporal Price Discrimination and Peak-Load Pricing

Intertemporal price discrimination: Practice of separating consumers with different demand functions into different groups by charging different prices at different points in time.

Consumers are divided into groups by changing the price over time. Initially, the price is high. The firm captures surplus from consumers who have a high demand for the good and who are unwilling to wait to buy it. Later the price is reduced to appeal to the mass market.

Example 11.4 (How to Price a Best-selling Novel) Publishing both hardbound and paperback editions of a book allows publishers to price discriminate.

Some consumers want to buy a new bestseller as soon as it is released, even if the price is $25. Other consumers, however, will wait a year until the book is available in paperback for $10.

The key is to divide consumers into two groups, so that those who are willing to pay a high price do so and only those unwilling to pay a high price wait and buy the paperback.

It is clear, however, that those consumers willing to wait for the paperback edition have demands that are far more elastic than those of bibliophiles. It is not surprising, then, that paperback editions sell for so much less than hardbacks.

Peak-load pricing: Practice of charging higher prices during peak periods when capacity constraints cause marginal costs to be high.

Demands for some goods and services increase sharply during particular times of the day or year.

Charging a higher price during the peak periods is more profitable for the firm than charging a single price at all times.

It is also more efficient because marginal cost is higher during peak periods.

11.7 Bundling

Bundling: Practice of selling two or more products as a package.

To see how a film company can use customer heterogeneity to its advantage, suppose that there are two movie theaters and that their reservation prices for the two films are as follows:

	Gone with the Wind	Getting Gertie's Garter
Theater A	$12,000	$3,000
Theater B	$10,000	$4,000

Table 11.4: Selling two films as a package

If the films are rented separately, the maximum price that could be charged for Wind is $10,000 because charging more would exclude Theater B. Similarly, the maximum price that could be charged for Gertie is $3000. The total revenue is $26,000.

But suppose the films are bundled. Theater A values the pair of films at $15,000 ($12,000 + $3000), and Theater B values the pair at $14,000 ($10,000 + $4000). Therefore, the film company can charge each theater $14,000 for the pair of films and earn a total revenue of $28,000.

Why is bundling more profitable than selling the films separately? Because the relative valuations of the two films are reversed. The demands are negatively correlated-the customer willing to pay the most for Wind is willing to pay the least for Gertie.

Suppose demands were positively correlated-that is, Theater A would pay more for both films:

	Gone with the Wind	Getting Gertie's Garter
Theater A	$12,000	$4,000
Theater B	$10,000	$3,000

Table 11.5: Selling two films as a package

If we bundled the films, the maximum price that could be charged for the package is \$13,000, yielding a total revenue of \$26,000, the same as by renting the films separately.

Chapter 12

Monopolistic Competition and Oligopoly

Competition and monopoly lie at opposite ends of the market spectrum. While competition is characterized by many firms, unrestricted entry, and homogeneous products, a monopoly is the sole producer of a product with no close substitutes. Falling between competition and monopoly are two other types of market structures: monopolistic competition and oligopoly, which describe the major part of remaining firms. In substance, ***monopolistic competition is closer to competition; it has many firms and unrestricted entry or exit***, like the competitive model, but its products are differentiated. Oligopoly, on the other hand, is more like monopoly; ***it is characterized by a small number of large firms*** producing either a homogeneous product like steel, or differentiated products like automobiles, and acting to each other in their actions. In the following, we will examine some of these models, noting the similarities as well as the differences between these models.

12.1 Monopolistic Competition

The Characteristics of Monopolistic Competition

- Many firms.
- Firms compete by selling differentiated products that are highly substitutable for one another but not perfect substitutes; in other words, each firm's product is slightly different from other firms in the industry so that the cross-price elasticities of demand are large but not infinite. It has only limited market power.
- Free entry and exit: it is relatively easy for new firms to enter the market with their own brands and for existing firms to leave if their products become unprofitable.
- Firms engage in nonprice competition — advertising is important.

Examples: Soft drinks; Wine; Chinese restaurants in NYC.

12.1.1 Short-Run Equilibrium

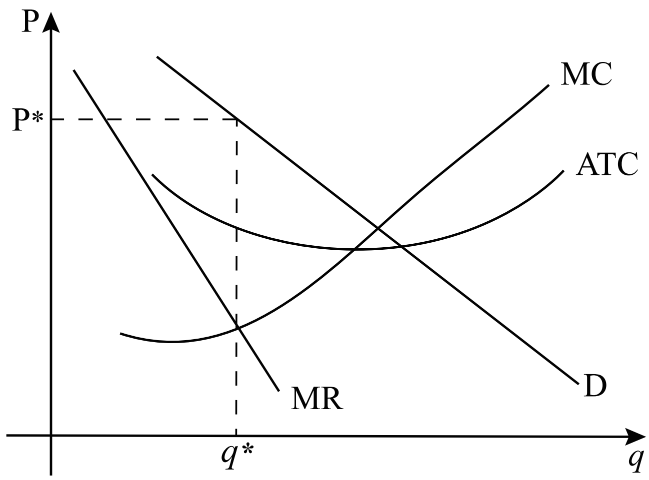

Figure 12.1 Short-run profit maximization under monopolistic competition

The same as monopolist. Because the firm is the only producer of its brand, it faces a downward-sloping demand curve. Price exceeds marginal cost and the firm has monopoly power. E.g., profit maximization.

12.1.2 Long-Run Equilibrium

Recall that a monopolist can earn positive economic profits in the long run. The above equilibrium could represent a monopolist's LR equilibrium. However, this does not hold for a monopolistically competitive firm as there are no barriers to entry.

Entry Eliminates Profits:

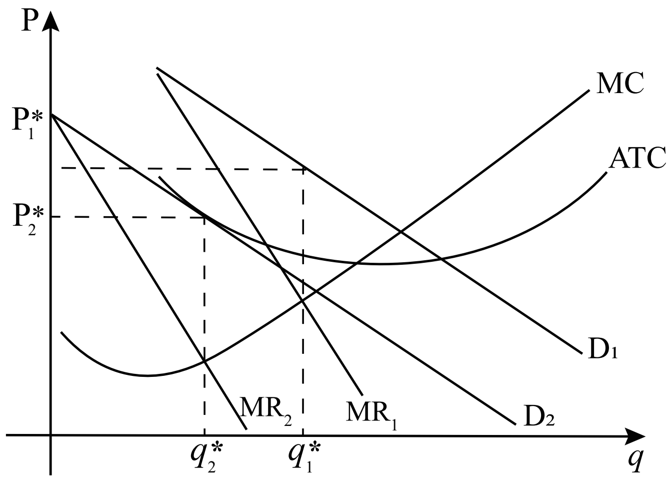

Figure 12.2 Entry eliminates profits under monopolistic competition

In the long run firms will enter the industry. Industry demand must be divided between more firms ⇒ demand for each firm is reduced. (D and MR shift to the left until profits are normal as depicted in Figure 12.2.) The long-run equilibrium is given by (p_2, q_2^*).

Exit Eliminates Losses:

As firms exit, the market demand is spread over fewer firms ⇒ demand for each remaining firm's output increases ⇒ D & MR shift to the right as depicted in Figure 12.3.

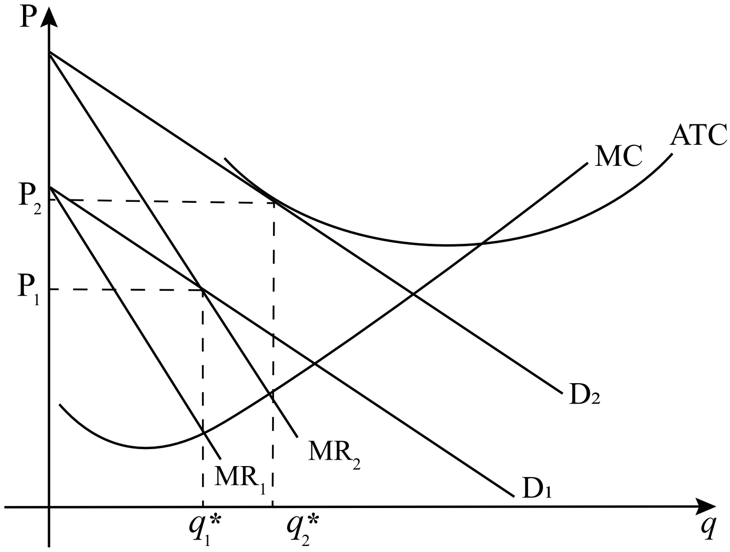

Figure 12.3 Exit eliminates losses under monopolistic competition

12.1.3 Monopolistic Competition and Economic Efficiency

Under perfect competition, price equals marginal cost. The demand curve facing the firm is horizontal, so ***the zero-profit point occurs at the point of minimum average cost***.

However, under monopolistic competition, the demand curve is downward-sloping, so ***the zero profit point is to the left of the point of minimum average cost***. Thus, price exceeds marginal cost. Thus ***there is a deadweight loss***, the same as the monopoly.

Is monopolistic competition then a socially undesirable market structure that should be regulated? The answer — for two reasons — is probably no:

(1) In most monopolistically competitive markets, **monopoly power is small**. Usually enough firms compete, with brand-

s that are sufficiently substitutable, so that no single firm has a big monopoly power. Any resulting deadweight loss will therefore be small. And because firms' demand curves will be fairly elastic, average cost will be close to the minimum.

(2) Any inefficiency must be balanced against an important benefit from monopolistic competition: **product diversity**. Most consumers value the ability to choose among a wide variety of competing products and brands that differ in various ways. The gains from product diversity can be large and may easily outweigh the inefficiency costs resulting from downward-sloping demand curves.

Example 12.1 (Monopolistic Competition for Soft Drinks and Coffee) The markets for soft drinks and coffee illustrate the characteristics of monopolistic competition. Each market has a variety of brands that differ slightly but are close substitutes for one another.

	Brand	Elasticity of Demand
Colas:		
	RC Cola	-2.4
	Coke	-5.2 to -5.7
Grand Coffee:		
	Folgers	-6.4
	Maxell house	-8.2
	Chock Full o' Nuts	-3.6

Table 12.1: Elasticity of demand for Colas and Coffee

With the exception of RC Cola and Chock Full o' Nuts, all the colas and coffees are quite price elastic. With elasticities on the order of -4 to -8, each brand has only limited monopoly power. This is typical of monopolistic competition.

12.2 Oligopoly

Oligopoly is an industry characterized by a few large firms producing most of or all of the output of some product, and reacting to each other in pricing and output decisions.

Characteristics of Oligopoly:

a) **Economies of scale**: It only takes a few firms of size q to supply the whole market. Scale economies may make it unprofitable for more than a few firms to coexist in the market; patents or access to a technology may exclude potential competitors; and the need to spend money for name recognition and market reputation may discourage entry by new firms. These are "natural" entry barriers — they are basic to the structure of the particular market. In addition, incumbent firms may take strategic actions to deter entry. Managing an oligopolistic firm is complicated because pricing, output, advertising, and investment decisions involve important strategic considerations, which can be highly complex.

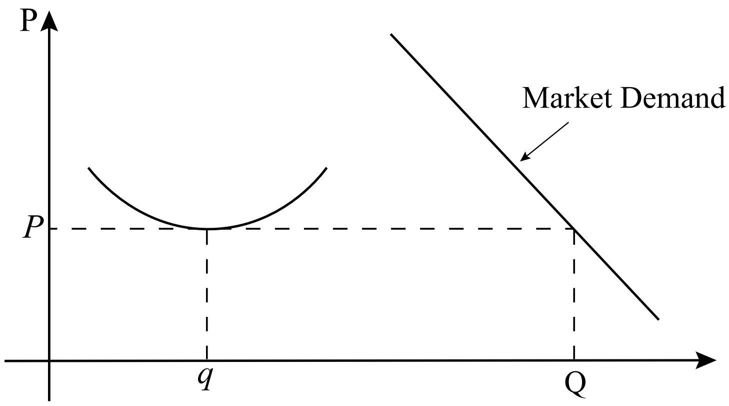

Figure 12.4 Downward-sloping demand curve under oligopoly

b) **Mutual interdependence among firms** - since there are only a few firms in the market each firm must react to other firms′ actions.

c) **Nonprice competition and price rigidity**. Price war is last alternative (fear of lowering profits). Competition relies on advertising and product differentiation.

d) **Temptation for firms to collude in setting prices.** Firms may want to maximize collective profits. This is illegal in the U.S.

e) **Incentive for firms to merge**. There is "perfect collusion" when the industry becomes a monopoly.

f) **Substantial barriers to entry** - such as

(i) **economies of scale**

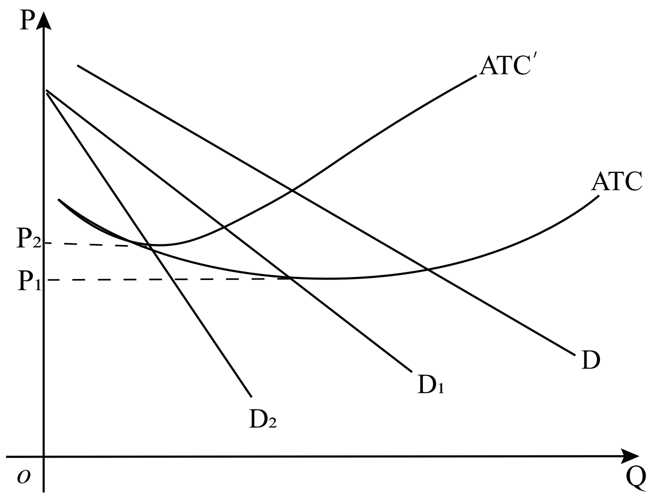

Figure 12.5 Individual demand and market demand under oligopoly

- D: demand faced by the first firm with no competitors.
- D_1: demand faced by the first firm with one competitor.
- D_2: demand faced by the competitor ($D = D_1 + D_2$).

- ATC: the same for both firms, with economies of scale.

Strategy for firm 1:

Charge a price below P_2 and above P_1. The potential loss of firm 2 will prevent it from entering the market.

Note: $ATC' \Rightarrow$ the 1st firm cannot use price to keep the 2nd firm out of the market.

(ii) **cost structure**

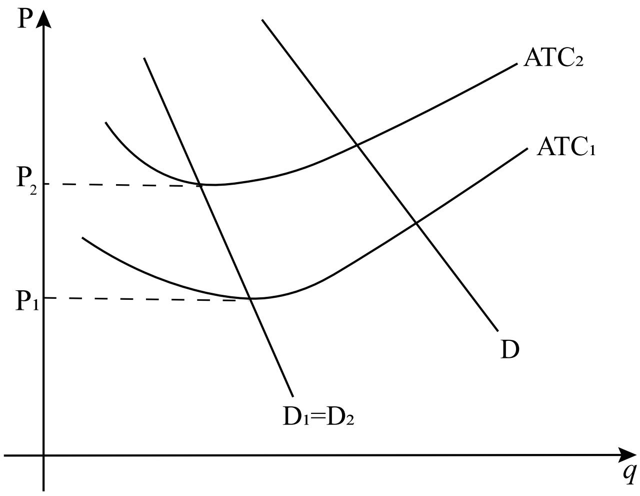

Figure 12.6 Oligopolistic firms with heterogeneous cost structures

Assume:

- firm 2 steals half of the market.
- costs for new firms are higher ($ATC_2 > ATC_1$).

Strategy for firm 1:

The first firm can keep the 2nd firm out by choosing a price between P_1 and P_2. P_2 is referred to as the limit price because any price greater than this will induce entry.

Therefore, in oligopolistic markets, the products may or may not be differentiated. What matters is account for most or all of total production. In some oligopolistic that only a few firms markets, some or all firms earn substantial profits over the long run because barriers to entry make it difficult or impossible for new firms to enter. Oligopoly is a prevalent form of market structure. Examples of oligopolistic industries include automobiles, steel, aluminum, petrochemicals, electrical equipment, and computers.

Strategic Interactions among Oligopolists

In an oligopolistic market, a firm sets price or output also based on strategic considerations regarding the behavior of its competitors. Then, game theory will paly in important role in studying the behavior of oligopolistic firms.

Game theory: Situation in which players (participants) make strategic decisions that take into account each other's actions and responses.

It can provide us with a way to examine the nature of interdependence among rival firms and to understand the role of uncertainty in their pricing and output decisions.

A game is described by the number of players, the set of strategies or actions, the payoffs of players, the forms of game (static or dynamic), and the information (complete or incomplete information) available to the players.

Payoff: Value associated with a possible outcome.

Strategy: Rule or plan of action for playing a game.

Optimal strategy: Strategy that maximizes a player's expected payoff.

It is essential to understand your opponent's point of view and to deduce his or her likely responses to your actions.

The basic and most common used solution concept of a game is Nash equilibrium, introduced by John Nash (1928-2015), a Nobel Laureate in Economics and Abel Prize winner, who inspired the movie "A beautiful

Mind".

Nash equilibrium: A profile of strategies or actions in which each firm does the best it can given its competitors' actions. In other words, in Nash Equilibrium, each firm is doing the best given what its competitors are doing, and thus has no incentive to deviate from the equilibrium strategy.

Therefore, at a Nash equilibrium, each firm will no longer adjust his own strategy given his rational expectation on the opponents' strategy profile. Thus, ***an important feature of Nash equilibrium is the consistency between belief and choice***. That is, a choice based on a belief is optimal, and the belief supporting this choice is correct.

A much stronger solution concept is dominant strategy equilibrium.

Dominant strategy equilibrium: Outcome of a game in which each firm is doing the best regardless of what its competitors are doing. That is, it is strategy-proof (no strategy will be played).

As long as there is dominant strategy equilibrium, firms will play it.

Duopoly: Market in which two firms compete with each other.

In the following we will examine common models that deal with oligopolistic behavior. As we discuss the models, keep in mind how different assumptions about rival behavior change the outcome.

12.3 Quantity Competition

12.3.1 Cournot Model

Cournot model: Oligopoly model, introduced by the French economist Cournot in 1838, in which firms produce a homogeneous product, each firm treats the output of its competitors as fixed, and all firms decide simultaneously how much to produce.

The model is then on quantity competition with homogeneous products, and shows how uncoordinated output decisions between rival firms could interact to produce an outcome that lies between the competitive and monopolistic equilibria.

For example, suppose there are two firms that produce identical products, marginal costs are zero, the market demand curve is known to both sellers and is linear, and each firm believes that its rival is insensitive to its own output.

Initially, firm A views D_T as its demand curve and produces $Q_1 = \frac{T}{2}$ where MC equals MR_T. Firm B enters the market and assumes firm A will continue to produce $Q_1 = \frac{T}{2}$ units so it sees D_B as its demand curve and produces $Q_2 = \frac{T-Q_1}{2}$. Firm A then readjusts its output to $Q_3 = \frac{T-Q_2}{2}$. The adjustment process continues until an equilibrium is reached with each firm producing $1/3$ of T.

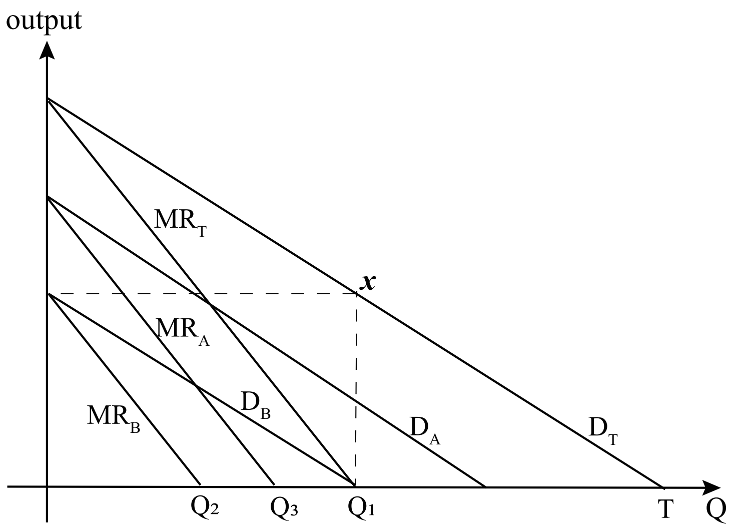

Figure 12.7 Market equilibrium under Cournot competition

Note that D_B is a residual demand curve which is obtained by subtracting the amount sold by firm A at all prices from the market $Q_2 = \frac{T-Q_1}{2}$, firm A will readjust its output and will view D_A as its new demand curve. D_A is a residual demand curve and is obtained subtracting firm B's output from D_T.

In the Cournot model, uncoordinated rival behavior produces a determinate equilibrium that is more than the monopoly output $Q_1 = \frac{T}{2}$, and two-thirds the competitive equilibrium output.

Reaction Curves and Cournot Equilibrium

Reaction curve: Relationship between a firm's profit-maximizing output and the amount it thinks its competitor will produce.

Firm 1's reaction curve shows how much it will produce as a function of how much it thinks Firm 2 will produce.

Firm 2's reaction curve shows its output as a function of how much it thinks Firm 1 will produce.

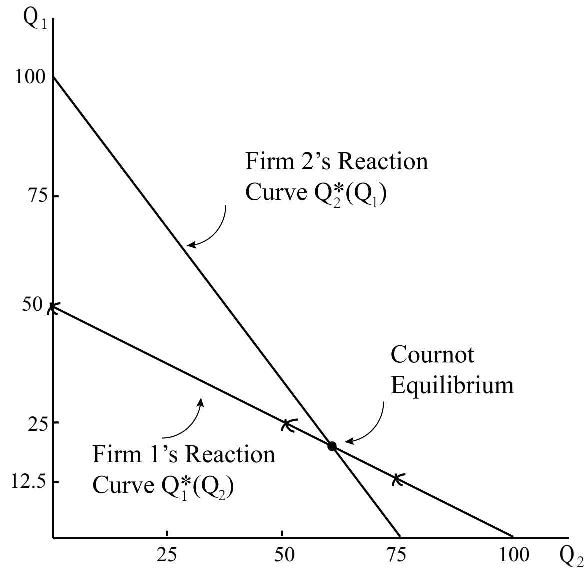

Figure 12.8 Reaction Curves and Cournot Equilibrium

Cournot equilibrium: Equilibrium in the Cournot model in which each firm ***correctly assumes*** how much its competitor will produce and sets its own production level accordingly. Cournot equilibrium is an example of a Nash equilibrium (and thus it is sometimes known as the Cournot-Nash equilibrium).

In a Nash equilibrium, each firm is doing the best it can ***given what its competitors are doing***. As a result, no firm would individually want to change its behavior. In the Cournot equilibrium, each firm is producing an amount that maximizes its profit ***given what its competitor is producing***, so neither would want to change its output.

Algebraically, suppose that the market demand function is $P(q_1 + q_2)$. Firm 1 wishes to find q_1^* which maximizes profits by taking firm 2's level of output as fixed.

$$\max_{q_1} \quad P(q_1 + \bar{q}_2)q_1 - c_1(q_1).$$

Similarly, firm 2's problem is to find q_2^*.

$$\max_{q_2} \quad P(\bar{q}_1 + q_2)q_2 - c_2(q_2).$$

Under what circumstances will the actions of two firms be consistent? They will be consistent when the choice each firm makes is compatible with the other firm's expectations. In other words, the choice based on a belief is rational (optimal), and the belief supporting this choice is correct, and thus $\bar{q}_1 = q_1^*$ and $\bar{q}_2 = q_2^*$.

Example 12.2 (The Linear Demand Curve) Two identical firms face the following market demand curve:

$$P = 45 - Q,$$

where $Q = Q_1 + Q_2$. Also,

$$MC_1 = MC_2 = 3.$$

Total revenue for the two firms:

$$R_1 = PQ_1 = (45 - Q)Q_1 = 45Q_1 - Q_1^2 - Q_2Q_1,$$

and

$$R_2 = PQ_1 = (45 - Q)Q_2 = 45Q_2 - Q_1Q_2 - Q_2^2.$$

Then,

$$MR_1 = 45 - 2Q_1 - Q_2,$$

$$MR_2 = 45 - Q_1 - 2Q_2.$$

Setting $MR_1 = MC_1$ and solving for Q_1, we obtain Firm 1's reaction curve:

$$Q_1 = 21 - \frac{1}{2}Q_2. \tag{12.1}$$

By the same calculation, we can get firm 2's reaction curve:

$$Q_2 = 21 - \frac{1}{2}Q_1. \tag{12.2}$$

Cournot equilibrium is:

$$Q_1 = Q_2 = 14.$$

Total quantity produce is

$$Q = Q_1 + Q_2 = 28.$$

If the two firms collude (like a monopolist), then the total profit-maximizing quantity is:

Total revenue for the two firms:

$$R = PQ = (45 - Q)Q = 45Q - Q^2,$$

then

$$MR = \Delta R/\Delta Q = 45 - 2Q$$

Setting $MR = MC$, we find that total profit is maximized at $Q = 21$. Then, $Q_1 + Q_2 = 21$ is the collusion curve.

If the firms agree to share profits equally, each will produce half of the total output:

$$Q_1 = Q_2 = 10.5,$$

which is less than the output produced as an oligopolistic firm.

12.3.2 The Stackelberg Model—First Mover Advantage

Stackelberg model: Oligopoly model, introduced by Stackelberg in 1934, in which one firm sets its output before other firms do.

There is an advantage to move the first. To see this, let us reconsider Example 12.2.

Example 12.3 (The Linear Demand Curve (continued)) Suppose Firm 1 sets its output first and then Firm 2, after observing Firm 1's output, makes its output decision. In setting output, Firm 1 must therefore consider how Firm 2 will react.

$$P = 42 - Q$$

and

$$MC_1 = MC_2 = 3.$$

Firm 2's reaction curve:

$$Q_2 = 21 - \frac{1}{2}Q_1. \tag{12.3}$$

Firm 1's total revenue is then given by

$$\begin{aligned} R_1 &= PQ_1 = (45 - Q)Q_1 \\ &= 45Q_1 - Q_1^2 - Q_1Q_2, \\ &= 45Q_1 - Q_1^2 - Q_1(21 - \frac{1}{2}Q_1) \\ &= 24Q_1 - Q_1^2 - \frac{1}{2}Q_1 \end{aligned} \tag{12.4}$$

$$MR_1 = 24 - Q_1.$$

Setting $MR_1 = MC_1$ gives $Q_1 = 21$, and $Q_2 = 10.05$.

We conclude that Firm 1 produces twice as much as Firm 2 and makes twice as much profit. Going first gives Firm 1 an advantage.

12.4 Price Competition

12.4.1 Price Competition with Homogeneous Products—The Bertrand Model

The Bertrand Model is about price competition with homogeneous products.

Bertrand model: Oligopoly model, proposed by the French economist Joseph Bertrand in 1883, in which firms produce a homogeneous good, each firm treats the price of its competitors as fixed, and all firms decide simultaneously what price to charge.

This model results in the same outcome as perfect competition even there are just two firms.

Let's return to the duopoly example.

$$P = 30 - Q,$$

where $Q = Q_1 + Q_2$, and

$$MC_1 = MC_2 = 3.$$

$Q_1 = Q_2 = 9$, and in Cournot equilibrium, the market price is \$12, so that each firm makes a profit of \$81.

Now suppose that these two duopolists compete by simultaneously choosing a price instead of a quantity.

The only possible Nash equilibrium in the Bertrand model results in both firms setting price equal to marginal cost: $P_1 = P_2 = \$3$ because no firm can obtain higher profit by unilaterally changing its pricing strategy. The industry output is 27 units. So each firm produces 13.5 units, and both firms earn zero profit.

In the Cournot model, because each firm produces only 9 units, the market price is \$12. Now the market price is \$3. In the Cournot model, each firm made a profit; in the Bertrand model, the firms price at marginal cost and make no profit.

12.4.2 Price Competition with Differentiated Products

Now suppose that each of two duopolists has fixed costs of \$20 but zero variable costs, and that they face the same demand curves:

$$Q_1 = 12 - 2P_1 + P_2,$$

$$Q_2 = 12 - 2P_2 + P_2.$$

Firm 1's profit is:

$$\begin{aligned}\Pi_1 &= P_1Q_1 - 20 = P_1(12 - 2P_1 + P_2) - 20 \\ &= 12P_1 - 2P_1^2 + P_1P_2 - 20.\end{aligned}$$

Similarly, we can figure out that firm 2's profit is given by

$$\Pi_2 = P_2Q_2 - 20 = 12P_2 - 2P_2^2 + P_1P_2 - 20.$$

The two firms's profit maximizing prices are

$$\Delta\Pi_1/\Delta P_1 = 12 - 4P_1 + P_2 = 0, \tag{12.5}$$

$$\Delta\Pi_2/\Delta P_2 = 12 - 4P_2 + P_1 = 0. \tag{12.6}$$

Solving the above two equations, we obtain the Nash equilibrium:

$$P_1 = P_2 = 4.$$

Firm 1's reaction curve:

$$P_1 = 3 + \frac{1}{4}P_2.$$

Firm 2's reaction curve:

$$P_2 = 3 + \frac{1}{4}P_1.$$

The intersection of these two curves gives the above Nash equilibrium.

12.5 Kinked Demand Curve Model

Each firm in an oligopolistic industry is aware that any price change it makes will affect the sales of other firms in the industry. Thus, D and MR curves must take into account the reactions of rivals.

Price rigidity: Characteristic of oligopolistic markets by which firms are reluctant to change prices even if costs or demands change.

Kinked demand curve model: Oligopoly model in which each firm faces a demand curve kinked at the currently prevailing price: at higher prices demand is more elastic, whereas at lower prices it is less elastic.

In order to have a kinked demand curve, it is assumed that oligopolists expect that ***rivals would match price cuts but not price hikes***.

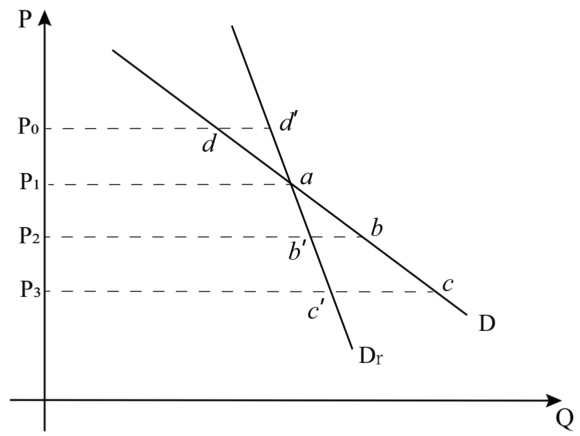

Figure 12.9 Demand curve which takes into account the reaction of rivals

As shown in Figure 12.9, we have:

- D: firm's demand curve when rivals do not react to price changes.

- D_r: firm's demand curve when rivals match any price change.

Along D, as P decreases, Q increases due to:

i) substitutability of this product for similar products in **other industries**; and
ii) purchases by customers who switch away from **other firms** in this industry.

Along demand curve D_r, we see that as P decreases, Q increases, but not by as much as that along demand curve D. This is because ii) above will no longer be true:

(1) without rival reactions

- P_1 decreases to $P_2 \Rightarrow a \to b$;
- P_1 decreases to $P_3 \Rightarrow a \to c$.

(2) with rival reactions

- P_1 decreases to $P_2 \Rightarrow a \to b'$ ($b' < b$);
- P_1 decreases to $P_3 \Rightarrow a \to c'$ ($c' < c$);
- If P_1 increases to P_0, no rival will increase its price. Original firm loses customers to other firms.

We therefore arrive at a **kinked demand curve**, as shown in Figure 12.10. That is, at higher prices demand is very elastic, whereas at lower prices it is inelastic.

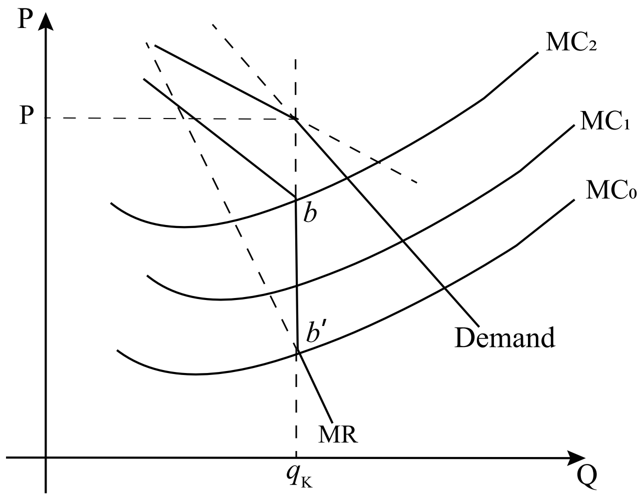

Figure 12.10 A kinked demand curve

To find MR here, we need to

i) identify q_K,
ii) draw MR for D until q_K is reached,
iii) extend D_r to vertical axis and draw MR beginning at q_K.

In the vertical segment b to b' the profit-maximizing rule $MR = MC$ yields the same (p, q) for any MC ($MC_2 \geq MC_1 \geq MC_0$). We thus come across the so-called **price rigidity**. Indeed, price rigidity is a characteristic of oligopolistic markets by which firms are reluctant to change prices even if costs or demands change.

12.6 Prisoners' Dilemma–Competition vs Collusion:

One game-theoretical model that has implications for oligopolist behavior is the so-called **prisoners' dilemma**. The prisoners' dilemma demonstrates how rivals could act to their mutual disadvantage. Suppose that two suspects, A and B, are apprehended and questioned separately about their involvement in crime. Without a confession, the district attorney has insufficient evidence for a conviction. The prisoners are unable to communicate with each other, and they are interrogated separately. During interrogation each is separately told the following outcomes about years in jail given in the "payoff matrix", a table showing payoff to each player given its decision and the decision of his competitor.

		Person B	
		confesses	doesn't confess
Person A	confesses	$(\underline{5}, \underline{5})$	$(\underline{1}, 10)$
	doesn't confess	$(10, \underline{1})$	$(2, 2)$

Table 12.2: Prisoners' Dilemma for Two Suspects.

Note that each "player" must individually choose a strategy (confess or not confess), but that the outcome of that choice depends on what the other player does.

In this setting, it is quite likely that players A and B will confess regardless what the opponent's choice is. That is, the strategy profile (**confess**, **confess**) is a dominant strategy equilibrium. Confessing makes sense since each prisoner is attempting to make the "best" of the "worst" outcomes.

It may be remarked that we can figure out an equilibrium for two-person games conveniently and quickly by using the following method. Consider the strategy of player 1, and for each strategy of player 2, find out the best responses of player 1. Draw a underline under its corresponding payoff. Similarly, find out the best responses of player 2 and underlines it. The strategy profile with both underlines is (**confess**, **confess**), which is a unique Nash equilibrium. The corresponding equilibrium payoff profile is (5, 5).

To see the relevance of the prisoners' dilemma to oligopoly theory, let us suppose that player A and player B are the firms in the same industry. The matrix in Table 12.3 identifies the payoffs about profits from an agreement to fix prices and share the market or from cheating on the collusive agreement. By the same reasoning as before, both are likely to cheat despite the fact that both will be worse off by cheating.

		player B	
		cheats	doesn't cheat
player A	cheats	$(\underline{10}, \underline{10})$	$(\underline{20}, 5)$
	doesn't cheat	$(5, \underline{20})$	$(15, 15)$

Table 12.3: Cooperation game of two oligopolistic firms. The cheating is a Nash equilibrium.

Example 12.4 (Price Competition (continued)) Consider the same example of duopoly: there are two firms, each of which has fixed costs of \$20 and zero variable costs. They face the same demand curves:

$$Q_1 = 12 - 2P_1 + P_2,$$

$$Q_2 = 12 - 2P_2 + P_2.$$

We found that in Nash equilibrium each firm will charge a price of \$4 and earn a profit of \$12, whereas if the firms collude, they will charge a price of \$6 and earn a profit of \$16, and whereas if Firm 1 keeps the colluding price \$6 and firm 2 cheats by setting price at \$4, their profits are

$$\Pi_1 = P_1 Q_1 - 20 = (6)[12 - (2)(6) + 4] - 20 = \$4$$

$$\Pi_2 = P_2 Q_2 - 20 = (4)[12 - (2)(4) + 6] - 20 = \$20,$$

respectively.

So if Firm 1 charges \$6 and Firm 2 charges only \$4, Firm 2's profit will increase to \$20. And it will do so at the expense of Firm 1's profit, which will fall to \$4.

Example 12.5 (Procter & Gamble in a Prisoner's Dilemma) The Procter & Gamble Company (P&G) is an American multinational consumer goods corporation headquartered in Cincinnati, Ohio, founded in 1837 by William

		Firm 2	
		charge \$4	charge \$6
Firm 1	charge \$4	$(\$\underline{12}, \$\underline{12})$	$(\$\underline{20}, \$4)$
	charge \$6	$(\$4, \$\underline{20})$	$(\$16, \$16)$

Table 12.4: **Payoff matrix**: Profits (or payoffs) to each firm given its decision and the decision of its competitor. The collusion is not a Nash equilibrium.

Procter and James Gamble. In 2014, P&G recorded \$83.1 billion in sales. It was renamed it as Procter & Gamble Health Limited in May 2019 after completing the acquisition of the consumer health division of Merck Groupin.

Candle maker William Procter, born in England, and soap maker James Gamble, born in Ireland, both emigrated to the US from the United Kingdom. They settled in Cincinnati, Ohio, initially and met when they married sisters Olivia and Elizabeth Norris. Alexander Norris, their father-in-law, persuaded them to become business partners, and in 1837 Procter & Gamble was created.

We argued that P&G should expect its competitors to charge a price of \$1.40 and should do the same. But P&G would be better off if it and its competitors all charged a price of \$1.50.

		Unilever and Kao	
		charge \$1.40	charge \$1.50
P& G	charge \$1.40	$(\$\underline{12}, \$\underline{12})$	$(\$\underline{29}, \$11)$
	charge \$1.50	$(\$3, \$\underline{21})$	$(\$20, \$20)$

Table 12.5: Payoff matrix for Pricing Problem.

Since these firms are in a prisoners' dilemma, it doesn't matter what Unilever and Kao do. P&G makes more money by charging \$1.40.

Advertising. Another way to illustrate the usefulness of the game-theoretical approach is to examine the interdependence of advertising decisions. Suppose that two firms are considering their advertising budgets.

They have two strategies, a large budget or small budget. The payoff matrix is the profits they get under different strategies. The resulting equilibrium strategy profile is (**large budget**, **large budget**).

		B	
		Small Budget	Large budget
A	Small Budget	$(10, 10)$	$(5, \underline{18})$
	Large budget	$(\underline{18}, 5)$	$(\underline{8}, \underline{8})$

Table 12.6: Payoff matrix for advertising game.

Unfortunately, not every game has a dominant strategy for each player

		B	
		Small Budget	Large budget
A	Small Budget	$(\underline{10}, \underline{8})$	$(6, 5)$
	Large budget	$(4, 2)$	$(\underline{8}, \underline{6})$

Table 12.7: Modified advertising game.

Now Firm A has no dominant strategy since its optimal decision depends on what Firm B does. If Firm B chooses a small budget for advertising, Firm A does best by also choosing a small budge for advertising; but if Firm B chooses a large budget for advertising, Firm A does best by also choosing a large budge for advertising. Notice that although it does not have a dominant strategy equilibrium, it has two Nash equilibria: (**small budget**, **small budget**) and (**large budget**, **large budget**).

Furthermore, does the prisoners' dilemma doom oligopolistic firms to aggressive competition and low profits? Not necessarily. Although our imaginary prisoners have only one opportunity to confess, most firms set output and price over and over again, continually observing their competitors' behavior and adjusting their own accordingly. This allows firms to develop reputations from which trust can arise. As a result, oligopolistic coordination and cooperation can sometimes prevail.

12.7 Dominant Firm Price Leadership Model

Price leadership: Pattern of pricing in which one firm regularly announces price changes that other (smaller) firms then match.

Price leadership model is another way to resolve the uncertainty of rivals' reactions to price changes. If one firm in the industry initiates a price change, and the rest of the firms traditionally follow the leader, there is no uncertainty about rival behavior.

A **dominant firm** is a firm with a large share of total sales that sets price to maximize its profit, taking into account the supply response of many smaller firms. **Price leadership** by the dominant firm occurs when the dominant firm in the industry sets a price that maximizes its profit and allows its smaller rivals to sell as much as they want at the price so set.

In some industries, a large firm might naturally emerge as a leader, with the other firms deciding that they are best off just matching the leader's prices, rather than trying to undercut the leader or each other. For example, we may consider the OPEC (Organization Petroleum Exporting Countries) as a dominant firm, and non-OPEC ones as smaller firms.

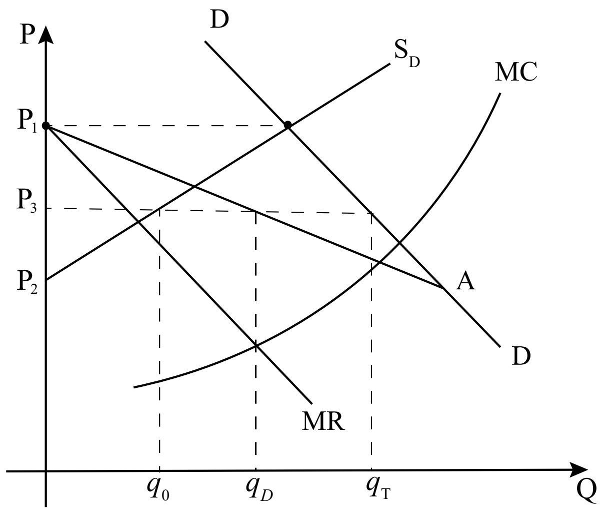

Figure 12.11 Equilibrium price and quantity under price leadership

Suppose the industry demand curve is DD. Since the other smaller firms in the industry will follow any price change initiated by the dominant firm, they become price takers, and adjust output until price equals their marginal cost. The residual demand curve confronting the dominant firm is obtained by subtracting the quantity the other smaller firms will produce shown by S_D from market demand, yielding P_1AD. The dominant firm produces q_D, where $MC_D = MR_D$, and charges P_3. Smaller firms produce q_0.

12.8 Cartels and Collusion

Cartel: Producers explicitly agree to cooperate in setting prices and output levels.

In all models discussed so far, the firms were assumed to behave independently. Each firm makes a specific conjecture regarding how other firms will respond to its action without any concern for how this affects the profits of the other firms. That is, there is no cooperation among firms.

The most important cooperative model of oligopoly is the cartel model. A **cartel** is an explicit agreement among independent producers to coordinate their decisions so each of them will earn monopoly profits. ***If producers adhere to the cartel's agreements and market demand is sufficiently inelastic, the cartel may drive prices well above competitive levels.***

Cartels are often international. While U.S. antitrust laws prohibit American companies from colluding, those of other countries are much weaker and are sometimes poorly enforced. Furthermore, nothing prevents countries, or companies owned or controlled by foreign governments, from forming cartels. For example, the OPEC cartel which is formed in 1960 is an international agreement among oil-producing countries which has succeeded in raising world oil prices above competitive levels.

Cartelization of a Competitive Industry

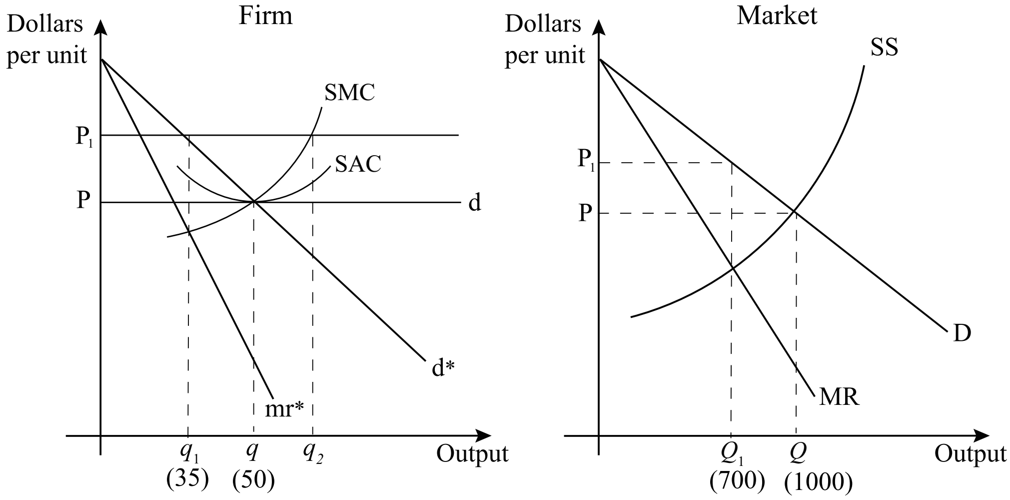

Figure 12.12 The emergence of cartels in a competitive market

Let us see how a group of firms in a competitive market can earn monopoly profit by coordinating their activities. We assume that the industry is initially in a long-run equilibrium, and then will identify the short-run adjustments (with existing plants) that the industry's firms can make to reap monopoly profits for themselves. Figure 12.12 shows this possibility.

Under competitive conditions, industry output is Q and price is P. If the firms in the industry form a cartel, output is restricted to Q_1 in order to charge price P_1, the monopoly outcome. Each firm produces q_1 and makes a profit at price P_1.

Firms can always make a larger profit by colluding rather than by competing. Acting independently, competitive firms are unable to raise price by restricting output, but when they act jointly to limit the amount supplied, price will increase. Cartel pricing can be analyzed by using the dominant firm model (see Figure 12.11).

There are two conditions for cartel success:

- First, a stable cartel organization must be formed whose members agree on price and output levels and then adhere to that agreement.

- The second condition, and may be the most important, is the potential for monopoly power. Even if a cartel can solve its organizational problems, there will be little room to raise price if it faces a highly elastic demand curve.

Why Cartels Fail?

If cartels are profitable for the members, why are not there many more? One reason is that in the United States they are illegal. But they were rare and were short-lived even before there were such laws. Three important factors appear to contribute to cartel instability.

1. Each firm has strong incentive to cheat on the cartel agreement as shown in prisoners' dilemma. (Say, the above figure, if one firm enlarge its output with price P_1, it can earn much more profit.) Yet, if many firms do so, industry output will increase significantly, and price will fall below the monopoly level. It is in each firm's interest to have other firms restrict their output while to increase its own output.
2. Members of the cartels will disagree over appreciate cartel policy regarding prices, output, market shares, and profit sharing. This is true when cost, technology, and size of firms are different.
3. Profit of the cartel members will encourage entry into the industry. If the cartel achieves economic profit its by raising the price, new firms have an incentive to enter the market. If the cartel cannot block entry of new firms, price will be driven back down to the competitive level as production from the "outsiders" reaches the market.

Chapter 14

Employment and Pricing of Inputs

The emphasis now shifts from product markets to input (or factor) markets. We will begin to look more closely at factors that determine the level of employment and prices of inputs used to produce a product. Firms are suppliers in product markets, and they are also demanders in input markets. Households and individuals are the demanders in product markets while the suppliers in input markets.

In this chapter, we will discuss the basic principles common to all input-market analyses, regardless of whether the input refers to labor, capital, or raw materials.

14.1 The Input Demand of a Competitive Firm

14.1.1 Firm's Input Demand: One Variable Input

We first introduce two concepts:

- **Derived input demand**: Demand for an input that depends on both the firm's level of output and the cost of inputs, and is derived from profit maximization.

- **Marginal revenue product (MRP)** of an input: Additional value resulting from the sale of output created by the use of one additional unit of an input.

Suppose that only one input (labor) is allowed to vary and the others are fixed. This is a short-run setting in which labor is the only variable input. By the law of diminishing marginal returns, the labor's marginal product (MP_L) curve slopes downward beyond some point.

The **the marginal revenue product of labor (MRP_L)** is then defined as

$$MRP_L = MR \times MP_L, \tag{14.1}$$

whether or not the output market is competitive.

However, in a competitive output market, $MR = P$, and then MRP_L reduces to the so-called the **marginal value product of labor** (MVP_L):

$$MVP_L = P \times MP_L, \tag{14.2}$$

which is downward slopping because the marginal product of labor falls as hours of work increase.

What is the optimal employment number of workers? The firm will hire up to the point where the marginal value product of labor is just equal to its marginal cost. In other words, the profit-maximizing firm will hire units of labor at the point where the marginal value product of labor is equal to the wage rate, i.e.,

$$w = MVP_L,$$

which gives us the ***the competitive firm's input demand curve for labor***.

Suppose that the daily wage rate is \$30 per worker. Then, the profit-maximizing labor-input units are $L^* = 20$. Therefore, factor markets are similar to output markets in many ways.

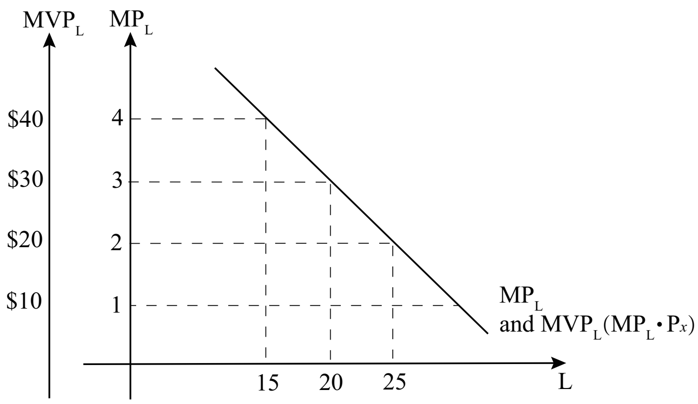

Figure 14.1 The input demand curve with only labor input variable, which is given by $MVP_L = MP_L \times P_x$ with $P_x = \$10$ for product x.

In fact, the factor market profit-maximizing condition $w = MVP_L$ is equivalent to the output market condition that marginal revenue equals marginal cost. To see this, since $MVP_L = MP_L \times P$, dividing both sides of equation by the marginal product of labor, we have

$$\frac{w}{MP_L} = P.$$

Recall that we have $w/MP_L = MC$. Therefore, the above equation is equivalent to the condition of profit maximization (i.e., $MC = P$) in term of output.

14.1.2 Firm's Input Demand: All Inputs Variable

In general, a change in an input price will lead the firm to alter the demand of other inputs. What is the input demand curve when all inputs are variable?

Suppose that at initial equilibrium the hourly wage rate $= \$30$ and $L = 20$. The firm is operating at point A on MVP_L where $K = 10$. Now suppose wage rate decreases to $w = \$\ 20$. If k keep constant ($K = 10$), the firm will increase the employment of labor to $L = 25$. If capital increases to $K = 12$ (for more workers need more "tools"), the entire MVP_L curve shifts upward. This adjustment leads to a further increase in labor to point C. Connecting points A and C, we get the firm's input demand curve.

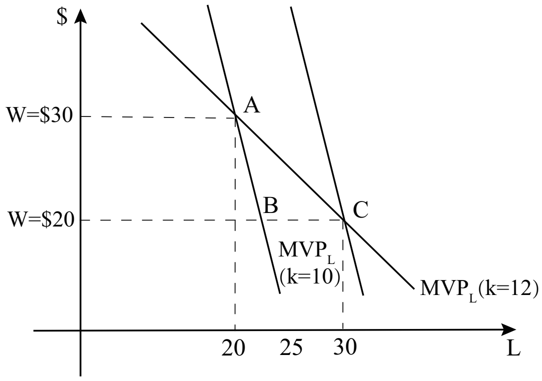

Figure 14.2 The input demand curve with all inputs variable.

14.2 The Input Demand of Competitive Industry

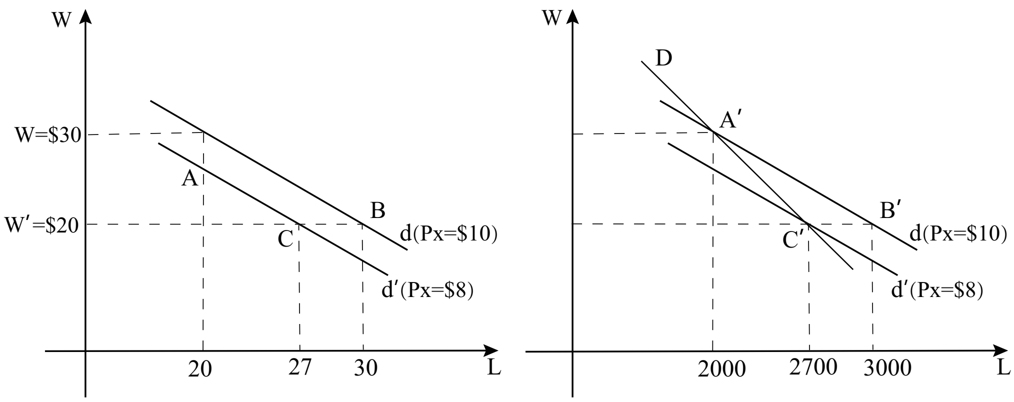

Figure 14.3 The industry demand curve in a competitive industry

The total quantity of an input hired by an industry is the sum of the quantities employed by the firms in the industry. To derive the industry input demand curve, we must therefore aggregate the input demand of all firms. However, when deriving a firm's input demand curve, we assume that the price of the product remains unchanged. This assumption is no longer true for an industry. When all firms simultaneously increase output, they can sell more output only at a lower price, which must be taken into account.

14.3 The Input Demand of Monopoly

A monopoly is defined as a firm that is the sole seller of some product, but a firm that has monopoly power in its output market does not necessarily have market power in its input markets.

Like a competitive firm, a monopoly makes decisions about input use such that marginal cost equals marginal revenue. Suppose input is labor. Then, for the profit-maximizing output of product x, we have

$$MC_L = MR_x.$$

Since $MC_L = \frac{w}{MP_L}$, we have

$$w = MP_L \times MR_x \equiv MRP_L.$$

Thus, the marginal revenue product curve MRP_L is the **monopolist's input demand curve**. As we know, MR_x of a monopoly must be lower than the price of output x, i.e., $MR_x < P$ at each level of output and at each level of employment of labor, so the MRP_L curve lies below the MVP_L curve which is the competitive demand curve.

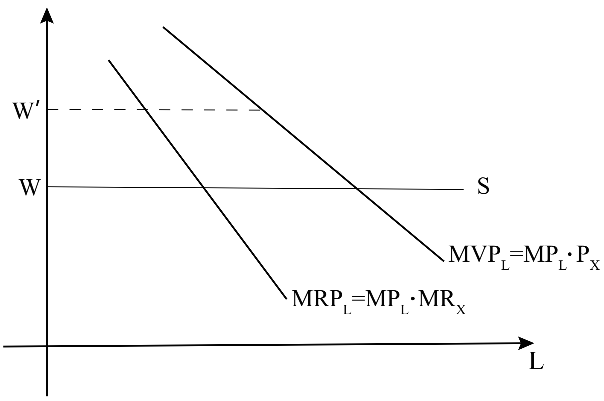

Figure 14.4 The relationship between MRP and MVP curves under monopoly

Note that the employment of labor, or any other input, is lower under monopolistic condition than under competitive condition. This result is consistent with the result we obtained previously: A monopoly produces less output than does a competitive industry.

14.4 The Supply of Inputs

The supply side of input markets deals with the quantities of inputs available at alternative prices.

Average expenditure curve (AV): Supply curve representing the price per unit that a firm pays for a good.

Marginal expenditure curve (ME): Curve describing the additional cost of purchasing one additional unit of a good.

Profit maximization requires that marginal revenue product be equal to marginal expenditure:

$$MRP = ME. \tag{14.3}$$

In the competitive case, since the price of the input equals marginal expenditure ($ME = w$), (14.3) reduces to

$$MRP = w. \tag{14.4}$$

The supply curve of inputs to ***all industries*** in the economy is almost vertical. For example, the total amount of labor can increase only if workers decide to work longer hours or if more people enter the labor force. Such responses to a higher wage rate may be so small.

Although the supply of input to all industries taken together may be vertical, this fact does not mean that the supply curve confronting any particular industry is vertical. The amount employed in a particular industry is subject to great variation. For example, if the wage rate paid to workers in the shoe industry should increase, workers in other industries would leave their jobs to go to work making shoes.

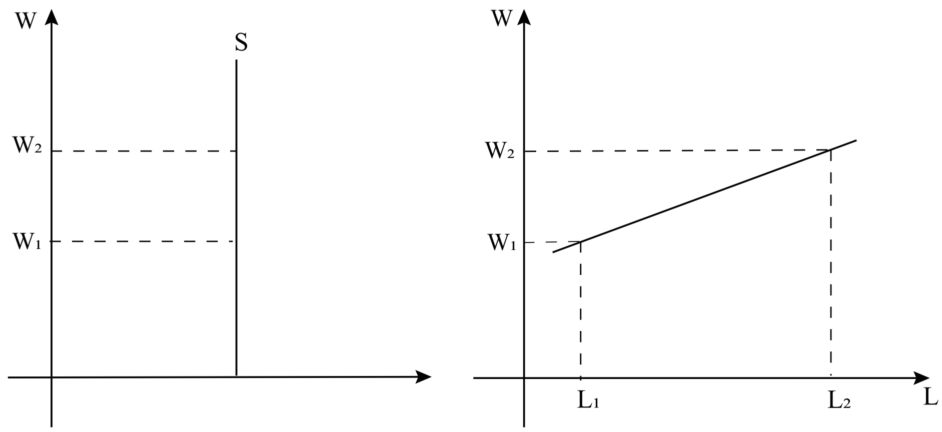

Figure 14.5 The diagram on the left is the supply curve for all industries, and the diagram on the right is the supply of an industry for labor input.

Because the shoe industry is only a small part of the entire labor market, its labor supply curve will be more elastic than the supply curve of labor for the entire economy. In fact, the supply curves of most inputs to most industries are likely to be either perfectly horizontal or slightly upward-sloping, as shown in Figure 14.5.

14.5 Labor Supply

Since consumers are supplers of labor, we can use the consumer theory to derive the labor supply curve.

14.5.1 The Income-Leisure Choice of a Worker

In our previous discussion of a consumer's choice, the consumer's income was assumed to be fixed. However, most people's income is not fixed but depends instead on, among other things, the decision about how much time the person will work. To investigate, we assume labor income is the only source of a worker's income, and the wage rate is fixed. Let

- $I =$ the weekly income;
- $L =$ the weekly leisure time;
- $H =$ the weekly working hours;
- $w =$ the wage rate per hour.

Then the income the consumer earns is $I = wH$. Since a week has $24 \times 7 = 168$ hours and $H + L = 168$, we have $I = w(168 - L)$, and thus

$$I + wL = 168w,$$

which is **the consumer's budget constraint**. The consumer's problem is choosing a combination of (I, L), such that he has highest utility.

Suppose the consumer's indifference curves are strictly convex over (I, L). We can find the optimal bundle by either graphical or mathematical approach. As usual, the equilibrium is the point of tangency between the budget line and an indifference curve.

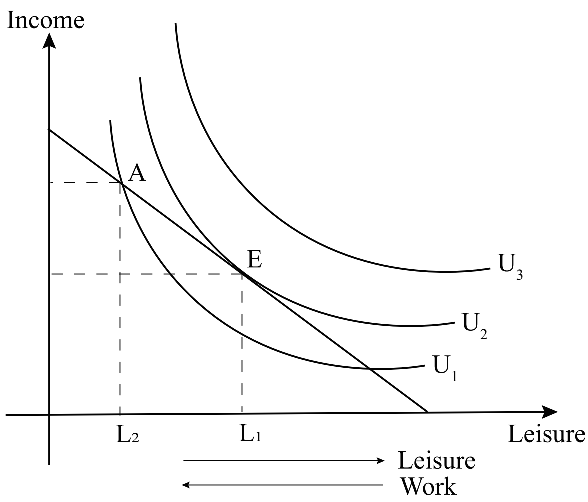

Figure 14.6 The income-leisure choice of a worker

14.5.2 The Supply of Working Hours

The above assumed that the wage rate was fixed. What happens if the wage rate changes? Will workers work longer hours at a higher wage rate? The answer depends on the consumer's preference.

The substitution effect dominating the income effect leads to more working hours

Suppose initial wage rate is $w_0 = \$10$, the optimal bundle is E. If the wage increases to $w_1 = \$15$, the new optimal bundle is E'. The substitution effect $= L_3 - L_1 < 0$. The substitution effect of a higher wage rate means to encourage a worker to have less leisure time, or to supply more hours of labor. The income effect $= L_2 - L_3 > 0$, which means the higher wage rate encourages the consumer to have more leisure time, or to supply less hours of labor since income and leisure are both normal goods. The total effect of the higher wage rate is the sum of the income and substitution effects. Although these effects operate in opposite directions, in this case the substitution effect is large, so the total effect is an increase in working hours from NL_1 to NL_2.

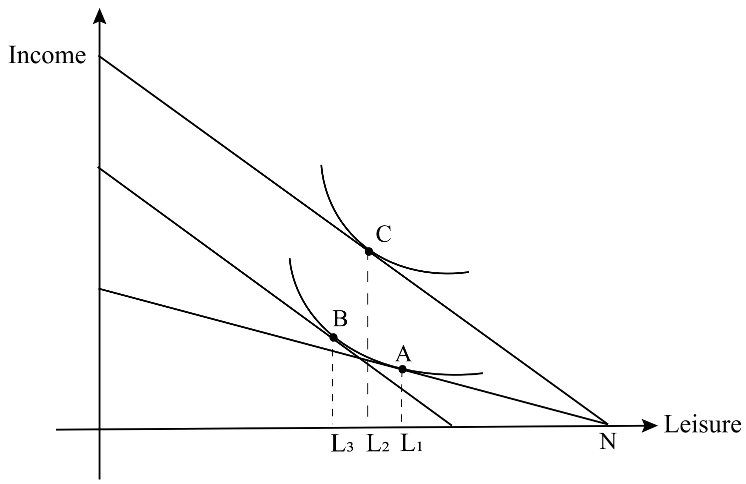

Figure 14.7 Consumer choice when substitution effect dominates income effect

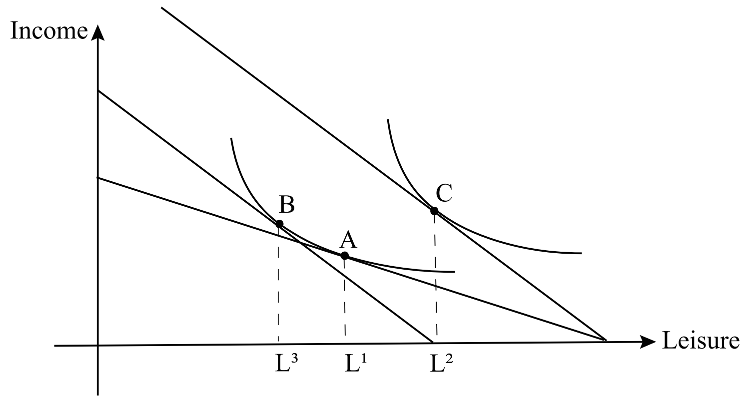

Figure 14.8 Consumer choice when income effect dominates substitution effect

The income effect dominating the substitution effect implies more leisure time (less working hours)

When the income effect is larger than the substitution effect, the conclusion is different from the above. The higher wage rate leads to a decrease in hours worked. Suppose that w further increases to $w = 20$. Total effect $= (L^2 - L^3) + (L^3 - L^1) = L^2 - L^1 > 0$, which means the consumer increases the leisure time.

14.5.3 Backward-bending Labor Supply Curve

When the wage rate increases, the hours of work supplied increase initially but can eventually decrease as individuals choose to enjoy more leisure and to work less. The backward-bending portion of the labor supply curve arises when the income effect of the higher wage (which encourages more leisure) is greater than the substitution effect (which encourages more work). That is, for low values of w, the substitution effect dominates. Beyond a threshold w^*, the income effect dominates.

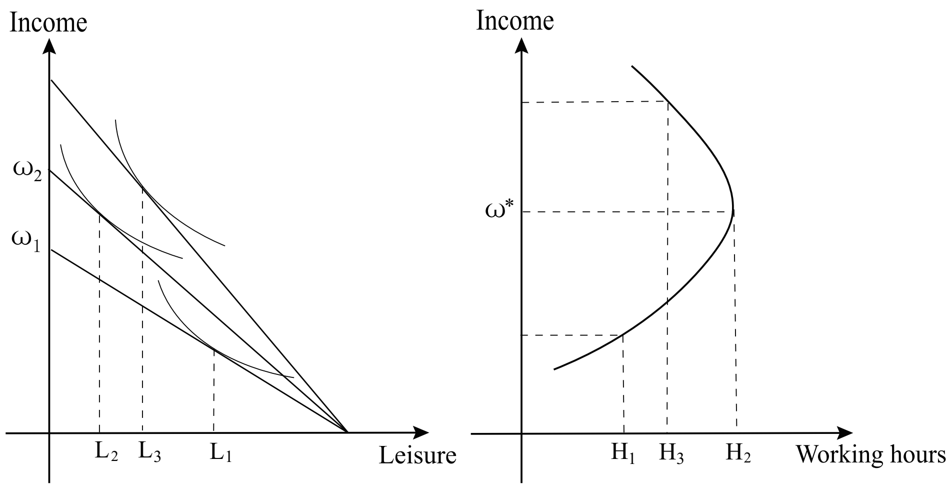

Figure 14.9 The emergency of backward-bending labor supply curve

Labor supply curve can be a vertical line.

Example 14.1 $U = L^\alpha I^\beta \Rightarrow MRS = \frac{\alpha I}{\beta L}$, where $\alpha > 0$, $\beta > 0$, and $\alpha + \beta = 1$.

By $MRS_{LI} = \frac{w}{1}$, we have $\frac{\alpha I}{\beta L} = w$ and thus

$$I = \beta w L / \alpha.$$

Substituting $I = \beta wL/\alpha$ into the budget constraint $I + wL = 168w$, we have

$$\frac{\beta wL}{\alpha} + wL = 168w$$

or

$$(\frac{\beta}{\alpha} + 1)wL = 168w,$$

and therefore

$$L^* = 168\alpha.$$

The consumer works 24α hours a day. This is true for all wage rate W. That is, L^* is independent of w.

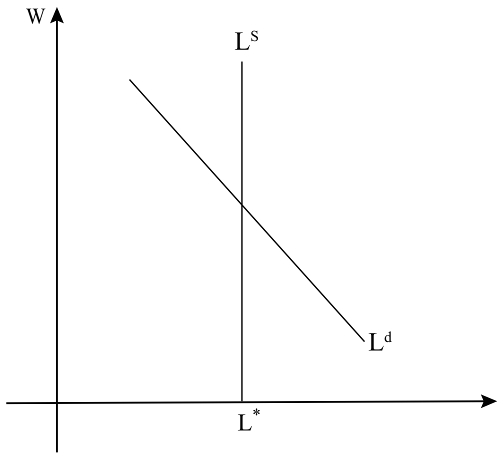

Figure 14.10 A vertical labor supply curve

The Market Labor Supply Curve

To go from an individual's supply curve of hours of work to the market supply curve, we need only add (horizontally sum) the responses of al- l workers competing in a given labor market. Thus, the market supply curve can also slope upward, bend backward, or be a vertical line.

14.5.4 The General Level of Wage Rate

We still use supply and demand analysis to investigate the level of wage rate. The supply curve of labor for this problem should indicate the total quantity of labor that will be supplied by all persons at various wage levels. The appropriate supply curve is the aggregate supply curve of hours of work discussed in the previous section.

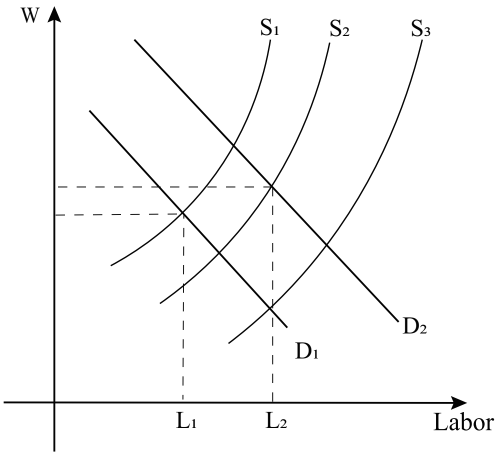

Figure 14.11 The demand and supply analysis of equilibrium wage rate

The aggregate demand curve for labor reflects the marginal productivity of labor to the economy as a whole. The aggregate demand curve for labor interacts with the aggregate supply curve to determine the general level of wage rates. Over time, both supply and demand increase. If demand increases faster than supply, wage rates tend to rise over time.

The productivity of labor is a main factor influencing the level of wages. This explains why real wage rates are so much higher in the U.S. than in the less developed countries. Marginal productivity is higher because of the factors that determine the position of the demand curve: capital, technology, and skill.

14.5.5 Why Wages Differ?

We know that there is a tendency for wage rates among firms or industries to equalize under the assumptions that workers were identical, and they evaluated the desirability of the jobs only in terms of the monetary wage rates. Dropping these assumptions leads to the conclusion that wage rates can differ among jobs and among people employed in the same line of work. Why is the wage rate for engineers higher than the wage rate for clerks? These differences are in full equilibrium with no tendency for the wage rates to equalize.

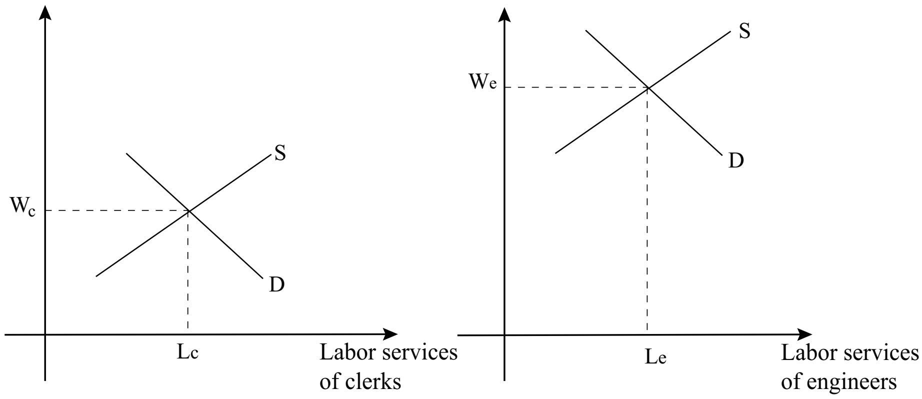

Figure 14.12 Wage rate gap between engineers and clerks

Factors leading to equilibrium wage differences

- **Equalizing wage differentials**: Workers currently employed as clerks may prefer their current jobs despite the financial difference. Monetary consideration is not the only factor, and sometimes not the most important factor, that influences the job selections of individuals. The differences of wages generated by preference on job choices are known as the **equalizing wage differentials** because the less attractive jobs must pay more to equalize the real advantages of employment among the jobs.
- **Differences in human capital investment**: Acquiring the skills to

become an engineer may have a significant cost. The wage for engineers may not be sufficiently high to compensate clerks for the training costs they would have to bear to become engineers.

- **Differences in ability**: Even if there were no training costs, clerks may not have the aptitude for science and mathematics necessary to work as engineers.

14.6 Industry Determination of Price and Employment of Inputs

14.6.1 The market equilibrium for labor input

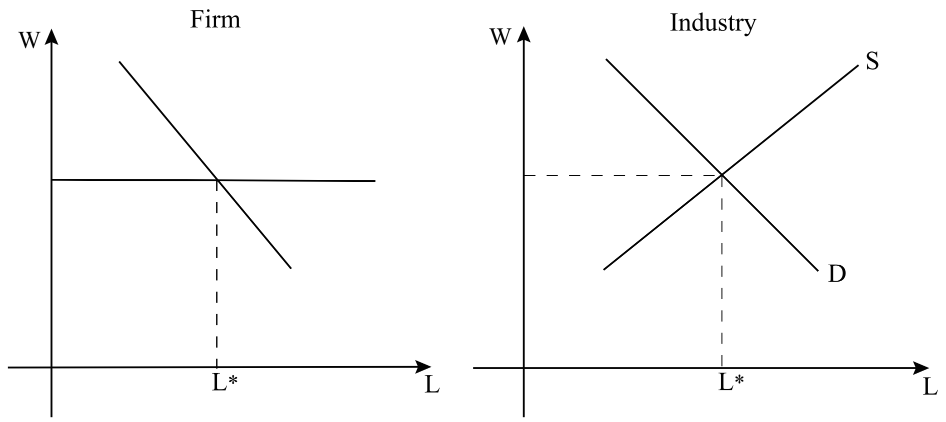

Figure 14.13 The market equilibrium for labor input

As usual, the market equilibrium of an input for a particular industry will be established when the quantity demanded equals the quantity supplied. Graphically, the equilibrium is shown by the intersection of the industry demand and supply curves. The position of a firm in equilibrium can be similarly determined.

Note that each firm is in the position of employing the quantity of the input at which the marginal value product equals the price.

Process of Input Price Equalization across Industries

When several industries employ the same input, the input tends to be allocated among industries so that its price is the same in every industry. If this were not true—if workers were receiving $40 in industry A and $30 in industry B—input owner would have incentive to shift inputs to industries where pay is higher, and this process tends to equalize input prices.

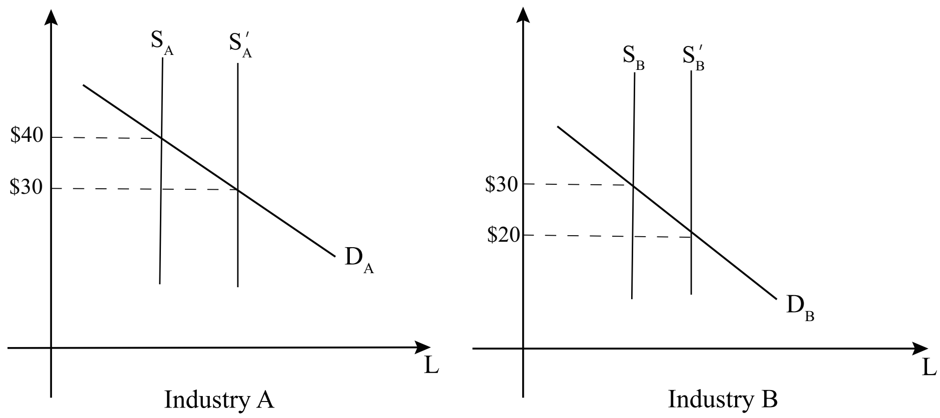

Figure 14.14 Wage rate equalization across two industries

14.6.2 Economic Rent

Rent: the payments made to lease the services of land, apartments, equipment, or some other durable assets.
Economic Rent: The difference between the payments made to a factor of production and the minimum amount that must be spent to obtain the use of that factor.

Economic Rent with Vertical Supply Curve

The vertical curve intercuts with the demand curve for the services to determine its price. The price and quantity are specified to indicate that we are not concerned with the sale price of the land but with the price for the services yielded by the use of land. When the supply curve of an input is vertical, the entire remuneration of the input represents ***economic rent*** since the same quantity would be available even at a zero price.

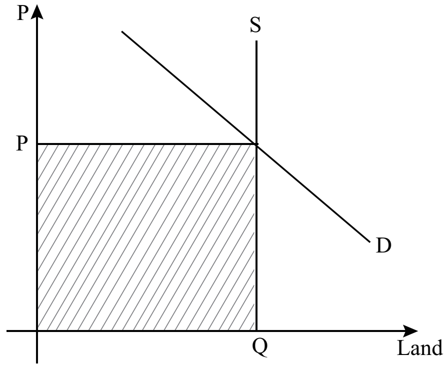

Figure 14.15 Economic rent under a vertical supply curve of land

Economic Rent with an Upward-sloping Supply Curve

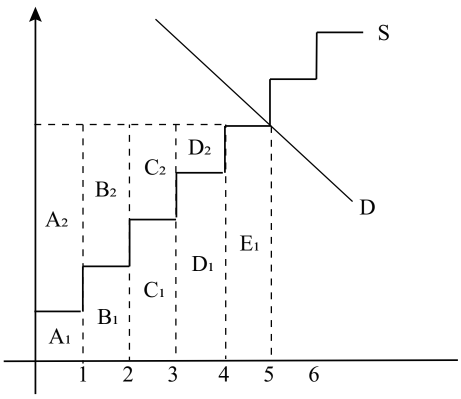

Figure 14.16 Economic rent with an upward-sloping supply curve

Consider the supply of college professors. With an upward-sloping input supply curve, part of the payment to input owners represents rent. In this case individuals A, B, C, and D receive rents equal to areas A_2, B_2, C_2, and D_2, respectively.

Whenever the supply curve slopes upward, part of the payments to inputs will be economic rents. The more inelastic the supply curve, the large rents are as a fraction of total payments. Note that rents are the net benefits received by owners of inputs from their current employment. They measure the gains from voluntary exchange.

14.7 Factor Markets with Monopsony Power

In some factor markets, buyers have **monopsony power** that allows them to affect the prices they pay. Often this happens either when one firm is a monopsony buyer or there are only a few buyers, in which case each firm has some monopsony power.

Firms with monopsony power buy large quantities and can negotiate lower prices than those charged smaller purchasers.

The **pure monopsony** here means a single firm that is the sole purchaser of some type of input. It faces the market supply curve of the input, a curve that is frequently upward sloping. An upward-sloping supply curve for labor means that the firm must pay a higher wage rate to increase the number of workers it employs. Thus, the marginal cost of hiring another worker, which is ME, is not equal to the wage rate it must pay to all workers ($ME > w$), and therefore ME lies above S.

The intersection of ME and the input demand curve determines employment, i.e.,

$$MVP_L = ME > w$$

if the firm is a competitor in its output market, and

$$MRP_L = ME > w$$

if the firm is a monopoly in its output market.

Thus, in comparison with competitive input conditions, employment is lower under monopsony and so is the wage rate paid. A similarity between monopsony and monopoly is apparently true from this conclusion. ***A monopoly restricts output in order to set a higher price while a***

monopsony restricts output (consequently less input) in order to pay a lower wage rate. A monopoly is able to charge a higher price because it faces a downward-sloping demand curve; a monopsony is able to pay a lower wage rate because it faces an upward-sloping supply curve.

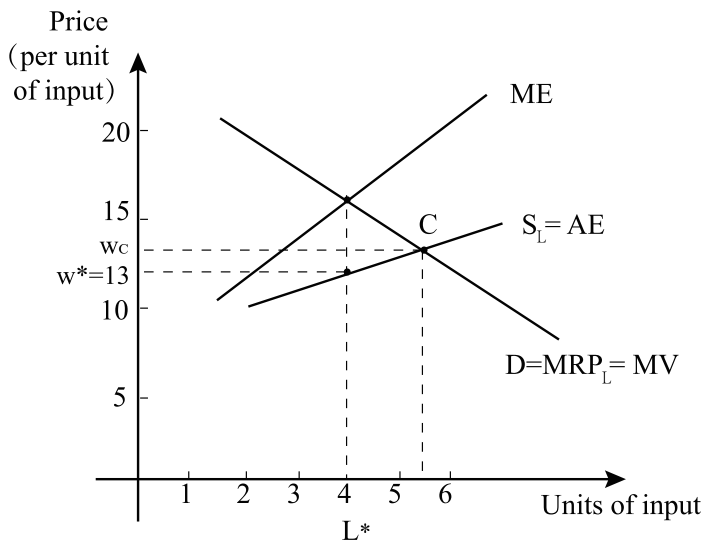

Figure 14.17 Marginal and average expenditure

Indeed, from Figure 14.17, when the buyer of an input has monopsony power, the marginal expenditure curve lies above the average expenditure curve because the average expenditure curve is upward sloping. The number of units of input purchased is given by L^*, at the intersection of the marginal revenue product and marginal expenditure curves. The corresponding wage rate w^* is lower than the competitive wage w_c.

14.8 Conclusions on Input Markets

Summarizing the discussions in this chapter, we can express the main results in one diagram depicted in Figure 14.18:

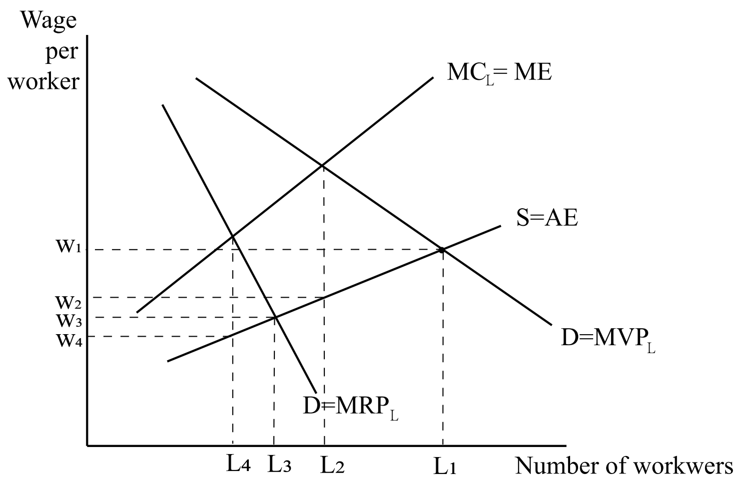

Figure 14.18 The profit-maximizing level of employment in various types of markets

(1) When the markets for labor and product are both perfectly competitive, the profit maximizing level of labor occurs where $MVP_L = w$. The number of units of input purchased is given by L_1, and the corresponding wage rate is w_1.

(2) When the market for labor has monopsony power and the market for product is perfectly competitive, the profit maximizing level of labor occurs where $MVP_L = ME > w$. The number of units of input purchased is L_2 that is lower than L_1, and the corresponding wage rate w_2 is lower than the competitive wage w_1.

(3) When the market for labor is perfectly competitive and the market for product has monopoly power, the profit maximizing level of labor occurs where $MRP_L = w$. The number of units of input purchased is L_3 which is lower than L_1, and the corresponding wage rate w_3 is lower than the competitive wage w_1.

(4) When the market for labor has monopsony power and the

market for product has monopoly power, the profit maximizing level of labor occurs where $MVP_L = ME > w$. The number of units of input purchased is given by L_4 lower than L_2, and the corresponding wage rate w_2 is lower than the competitive wage w_3.

Thanks for taking this course!