Economics and Tax Policy

Karl E. Case
Wellesley College

A Lincoln Institute of Land Policy Book

Published by
Oelgeschlager, Gunn & Hain
in association with the
Lincoln Institute of Land Policy

Copyright © 1986 by Oelgeschlager, Gunn & Hain, Publishers, Inc. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical (including photocopying and recording), or by any information storage or retrieval system, without the prior written consent of the publisher.

International Standard Book Number: 0-89946-209-x

Library of Congress Catalog Card Number: 85-29843

Printed in the U.S.A.

Oelgeschlager, Gunn & Hain, Publishers, Inc.
131 Clarendon Street
Boston, MA 02116 U.S.A.

Library of Congress Cataloging-in-Publication Data

Case, Karl E.
Economics and tax policy.

1. Economics. 2. Taxation. 3. Tax planning.
I. Title.
HB171.C374 1986 336.2 85-29843
ISBN 0-89946-209-x

Publication of a manuscript by the Lincoln Institute of Land Policy signifies that it is thought to be worthy of public consideration but does not imply endorsement of conclusions or recommendations. Views expressed in Lincoln Institute publications should be understood to be solely those of the authors and should not be attributed to the Lincoln Institute, its directors, officers, or staff members, or to the organizations that supported research leading to any such publication.

To Richard A. Musgrave in whose ideas and teachings germinated an enormous body of thought called public economics and a generation of economists dedicated to the theory and practice of building and preserving an efficient and just public sector governed by democratic processes.

Contents

List of Figures

List of Tables

Lincoln Institute Foreword

The Lincoln Institute of Land Policy is an educational institute dedicated to the development and exchange of ideas and information pertaining to land policy and property taxation. It is a school offering opportunities for instruction and research. It welcomes government officials, working practitioners, and students to the pursuit of advanced studies.

The Lincoln Institute is also a center for linking the university and the practice of government; for bringing together scholars, professionals, and officials; and for blending the theory and practice of land policy.

For more than a decade, the Lincoln Institute has jointly sponsored the Land Reform Training Institute (LRTI) in Taoyuan, Taiwan with the Council of Agriculture of the Republic of China. The LRTI is a training and research institute in the fields of land policy, land reform, and land taxation. The LRTI offers a variety of courses and research opportunities each year.

Professor Karl E. Case has taught at the LRTI, and as his course and others evolved, a need was recognized for a basic treatment of the linkage between economics and tax policy. Professor Case was asked to prepare such a manuscript. When he finished, it was realized that the text would be beneficial for all students of economics or property taxation, not just LRTI attendees.

We, therefore, include this book in our LILP/OG&H series with the recognition that it will help fulfill our mission of fostering the development and exchange of ideas and information pertaining to land policy and property taxation.

Frank Schnidman
Senior Fellow
April 1986

Acknowledgments

This book grew out of a course by the same name taught for many years in the International Tax Program at Harvard Law School. Oliver Oldman directs that program and was the first to suggest publishing some of the curricular materials. I am very grateful to the Lincoln Institute for Land Policy for supporting my work in the Tax Program and this book. Special thanks go to Sein Lin and Arlo Woolery.

For the last two years I have used this material teaching an abbreviated version of the same course at the Land Reform Training Institute (LRTI) in Taiwan. My students at Harvard and at the LRTI have been an enormous inspiration to me; they have taught me as much as I have taught them.

Maura Doyle drew all the diagrams and was a tireless research assistant. Catherine McKeen and Lori Martin turned out the manuscript in record time, and I am grateful for their long nights.

1

Introduction to Economics

What Is Economics and Why Study It?

An economy is a system or set of processes and institutions through which a society uses resources to meet human wants or needs. Every society has such a system, and in a very real sense, economy and society are one. Some would argue that the basic character of social life is determined by economic forces; certainly, individual and social behavior determine the character of an economy.

An understanding of basic economics is essential to the formulation and application of rational tax policy. Taxes change behavior, and those changes determine the ultimate effect of a tax. It is often difficult even to determine who ends up paying a tax. Consider, for example, the corporate income tax. The corporate income tax is imposed on an institution: the corporation. Institutions have no inherent tax-paying capacity; they are merely extensions of individuals. The question for rational tax policy is, who will end up bearing the burden of the corporate tax? There are three obvious candidates: (1) Owners of corporations may bear the tax through lower profits; (2) consumers of corporate products may face higher prices; (3) workers employed by corporations may find their wages reduced in the presence of the tax. The question is a simple one, and if the corporate tax is to be a major component of a revenue system, we should have an answer to it. This book provides tax policy makers with a knowledge of economics sufficient to formulate and begin to answer questions such as this one.

Resources and Production

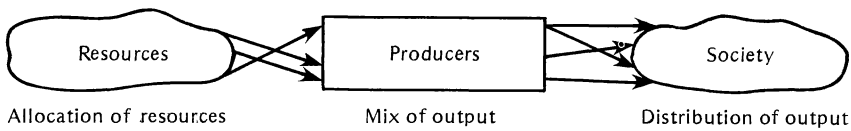
Every society, regardless of its political structure, has some system or mechanism for transforming resources into usable form. The term *re-*

sources is very broad. Some resources are the product of nature, such as land, minerals, energy, and timber. At any given time, the resources available to society also include some that have been produced, such as buildings and equipment, which are part of a nation's *capital stock* (a term that will be defined in more detail below). Human resources (labor and skills) are, of course, central. Time and knowledge may also be considered resources to be used by society.

Producers take resources and transform them into usable products or outputs. Many categories of activity fall under the heading of production. Private business firms purchase resources and produce products for the market. Governments purchase resources and produce things for society, too, such as national defense, a system of justice, police and fire protection, and sewer services.

Individual households often produce products for themselves. A household that owns its own home is, in essence, taking capital and land resources and producing *housing services* that it consumes itself. The Boston Symphony Orchestra is no less a producer than General Motors. The Boston Symphony takes capital (in the form of a building, musical instruments, lighting fixtures, and so forth) and combines it with land and labor to produce performances.

The economy is the mechanism that determines the allocation of resources among those engaged in production. It also determines the actual mix of outputs produced, and, ultimately, the final distribution. The economic mechanism determines *what* gets produced, *how* it gets produced, and *for whom* it gets produced. In studying economics, we focus on three outcomes: the *allocation* of resources among producers, the *mix* of output, and the *distribution* of that output among the members of society.



Positive Versus Normative Economics

In studying these various outcomes, economics asks two kinds of questions: positive and normative. Positive or descriptive economics attempts to understand the operation of the system in question without making value judgments about the outcomes. What determines the allocation of resources? What factors influence the mix of output? Why does society produce large automobiles and not small, fuel-efficient ones? Why

does one society produce many single-family homes, and another, small, attached housing units? What would happen if the corporate tax were abolished? Who would benefit? Who would lose?

Normative economics looks at the outcomes and asks, Are they good or bad? Should society alter the outcomes or change the system to make it better? Does the imposition of an income tax make the economic outcomes better or worse? Normative economics defines and uses a number of criteria for judging outcomes. Four criteria are applied frequently in economics:

1. Efficiency
2. Equity or fairness
3. Growth
4. Stability

Efficiency. An efficient change makes some people better off and no one worse off. However, nearly all changes that can be made in an economic system will leave some people better off and some worse off. If because of a change some individuals or groups gain more than the losers lose, the change is said to be *potentially efficient.* In theory, the gainers could compensate the losers for their losses, and something would be left over. The distinction between potential and actual efficient changes is generally ignored, and all such changes are simply called *efficient.*

There are many examples of efficient changes. When two people engage in voluntary exchange, both are made better off, and no one suffers a loss; a simple exchange is efficient. Some efficient changes are more complex. Several years ago in an economy move the Massachusetts Registry of Motor Vehicles substantially reduced the number of clerks in each office. Almost immediately, residents of the Commonwealth experienced long delays when registering their automobiles or obtaining driver's licenses. Standing in line uses time and energy that could otherwise be more productively used. A cost was clearly imposed on the citizens of the Commonwealth.

To calculate its value, one would have to ask, How much would citizens be willing to pay to avoid standing in those lines? In one office, it was estimated that approximately 500 people daily stood in line for about one hour each. If each person were willing to pay only $2 to avoid the line, the damage imposed is approximately $1,000 per day. If the registry were open 250 working days per year, the reduction in labor force at that registry office alone created a cost of $250,000 per year. It was also estimated that the savings realized by having fewer clerks at that office amounted to about $75,000 per year.

Clearly, if the government reinstated those employees, the result

would be efficient. The beneficiaries—those who would no longer have to stand in line to obtain registrations and driver's licenses—would gain significantly more than the costs imposed on the taxpayers of Massachusetts to finance the reinstatement.

It is often possible to estimate the potential efficiency gains or losses from larger changes in the economy. It will be argued later in this text that it is efficient to shift from a set of taxes that distorts economic decisions to a set of taxes that does not. Taking steps to reduce monopoly power is efficient. Imposing a tax on a firm that pollutes the air or water may be efficient if the tax is tied to the damage produced by the pollution. It can be shown that regulated prices in many instances are inefficient.

Equity. While the criterion of efficiency has a fairly precise definition, and one that can be applied with some degree of rigor, no two people agree on what is fair and unfair. To many, fairness implies a more equal distribution of income. Fairness may imply alleviating poverty, but the extent to which poverty should be reduced is hotly debated. Philosophers for centuries have wrestled with principles of justice to help guide social decisions.

Despite the lack of a precise definition, public policy-makers make judgments about fairness on a daily basis. Often, decisions made in the name of equity lead to inefficiency. During the oil embargo in 1973, the government imposed a ceiling on the price of gasoline at the pump. The argument was that it was unfair to sell gasoline only to the rich, who could afford it. Price restrictions such as this usually lead to an inefficient allocation of resources. Gasoline was no longer channeled to its most valuable uses.

Growth. As the result of technological change, the building of capital, and the acquisition of knowledge, societies learn how to produce products better. In the early days of the U.S. economy, it took nearly half the entire population to produce the food supply required. Today less than 3 percent of the population of the United States is engaged in agriculture, and that industry supplies a good portion of the world with food. When goods can be produced more cheaply, society can devote resources to the production of other things, and the standard of living improves. This increase in standard of living, which results from the development of new products and better production techniques, is called *economic growth.* Economic growth is a central concern of countries in the less-developed world.

Stability. Market economies are inherently unstable. For hundreds of years we have observed swings in economic performance that are

collectively referred to as *the business cycle.* Periods of expansion often lead to inflation, and boom periods inevitably end in slow-downs or recessions. Most economies place full employment and price stability high on their list of priorities. The 1982 recession in the United States left 12 million people unemployed and resulted in billions of dollars in lost output. It is not surprising that societies seek to avoid such losses.

The Methodology of Economics

Economics is a social science. It is not as elegant as physical science. It deals with regularities that emerge from the random social interaction of millions of people. To understand its basic underlying forces and tendencies, we must begin with abstraction. Critics of economics often focus on this abstraction.

To answer real-world questions one must deal explicitly with the institutional, social, and political context within which economic forces operate. It is impossible, however, to understand the operation of complex economic systems without exploring first the operation of simple ones. So economists build theories or models of simple economies stripped of institutional detail in order to expose basic economic forces. We study simple economies to understand more complex ones.

The Market

No two economic systems are identical; they vary enormously across societies. Some economic systems are characterized by central planning and extensive government involvement in production and distribution. Other economies are highly decentralized, with the allocation, mix, and distribution determined by the decisions of millions of individual firms and households. Common to all systems, however, is one institution: the market.

The market is simply an institution through which buyers and sellers interact and engage in exchange. The interactions among buyers and sellers may be very simple or enormously complex. In the Middle Ages, those who wished to engage in exchange simply brought their wares to a central place and traded. Today when a bank has extra cash, it seeks out other banks in need of reserves through the Federal Funds market. In that market, banks never communicate directly but only electronically through an intermediary in New York.

The next few chapters of this book explore the basic operation of markets in a simple economy. How markets behave determines the *allocation of resources,* the *mix of output,* and the *distribution of that output.*

2

Markets and Prices: The Basic Forces of Supply and Demand

Economic theory is, at its base, an attempt to analyze the way in which individuals and groups of individuals behave. Microeconomics focuses on decisions made by two fundamental decision-making units: households and firms.

Each household has a certain amount of income and wealth at its disposal and must decide how to divide it up between spending and saving and among the various goods and services available. The amount of income earned is likely to be affected by other decisions of the household: whether to obtain training, whether to enter the labor force or not, how much labor to supply, how to invest wealth that may have been inherited or saved.

Firms are individuals or groups of individuals that join forces to produce goods and services. Firms take resources purchased in markets and combine them in a process called *production* to produce goods and services desired by households. Firms must decide what to produce and how to produce it.

Firms and households relate to each other in two basic arenas: input markets and output markets (see Figure 2.1). Business firms produce and *supply* in output markets. Households viewing the range of goods available, incomes, and prices must decide how much to buy. Households *demand* in output markets. Inputs into the production process such as land and labor are sold by households to business firms. Firms *demand* in input markets, and households *supply*.

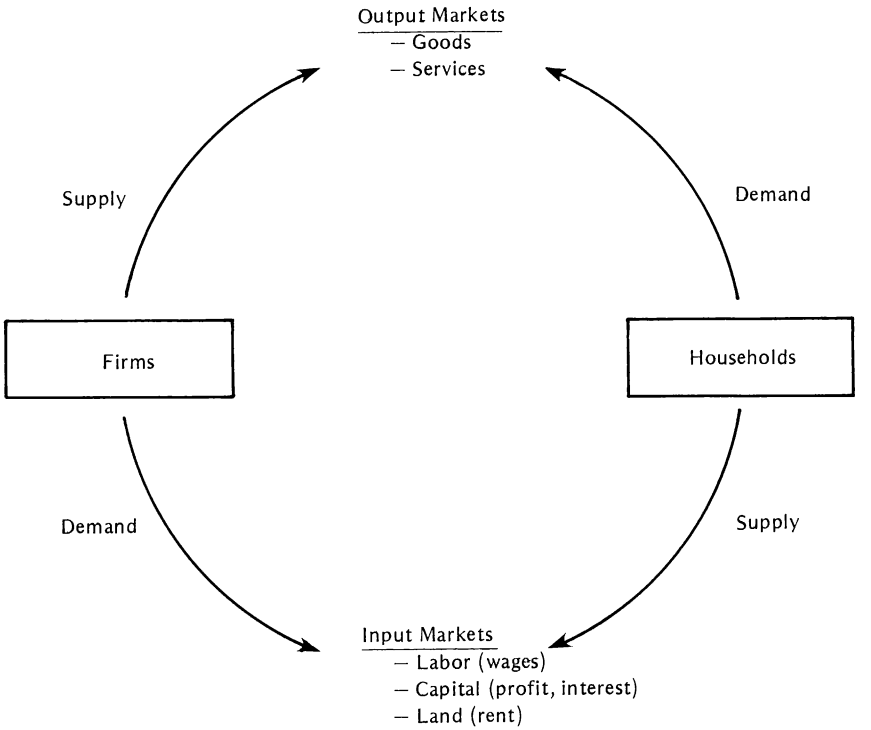


Figure 2.1. *The market allocation mechanism.*

Input markets and output markets are closely connected, and it is their combined operation that ends up determining the distribution of output, the allocation of resources, and the mix of output in a market system. This chapter explores the basic elements of market operation.

Household Demand in Output Markets

While households make a number of decisions simultaneously, we focus first on the amount that an individual might decide to purchase or *demand* of a single commodity in an output market in some given period of time, such as a month or a year. The amount of an individual commodity, X, which a household decides to demand might be expected to depend on a number of factors:

1. The income available to the household;
2. The amount of accumulated wealth of the household;

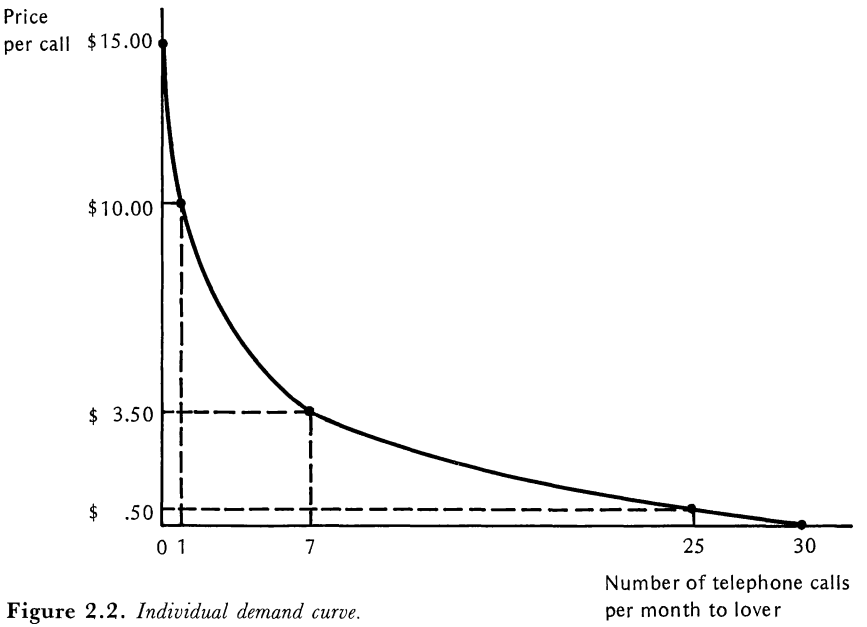


Figure 2.2. *Individual demand curve.*

3. The price of the product in question;
4. The prices of other products available to the household;
5. The tastes or preferences of the household.

The way in which we phrase the question is very important. We ask, How much would a household buy of product X if it could buy all it wants at the current market price? It is clear that the amount of product actually purchased by a household depends on the amount of product actually produced in the market. But demand at any moment may exceed or be less than supply. These differences are important, and it is essential to separate the supply decision from the demand decision in analyzing markets. Thus the phrase "if it could buy all it wants" is important.

We begin by focusing on the relationship between quantity demanded by an individual household and the price of the commodity or product in question *ceteris paribus*. The phrase *ceteris paribus* means "all else being equal." The demand curve is a graphical representation of the amount that a household would purchase at each individual price assuming that preferences, incomes, and other factors remain unchanged.

Figure 2.2 presents a hypothetical demand curve for a student who came to the United States to study economics leaving her lover behind. There are several things to notice about the curve. First, it intersects the Y, or price, axis, which means that there is a price above

which no calls will be made. In this case, when the price reaches $15 per call, the student simply writes more letters. Letters are a substitute for telephone calls, and they have a much lower price. As long as households have limited incomes or wealth, demand curves will hit the price axis. For any commodity, there is a price above which a household will not or cannot pay.

Second, the demand curve intersects the X, or quantity, axis. This suggests that even at a zero price there is some limited amount of the good in question that the household will consume. In this case, even if telephone calls cost nothing, the person would call thirty times per month and not more. Third, the demand curve has a negative slope. That means that a decline (negative change) in price seems to lead to an increase (positive change) in the quantity demanded; that is, our graduate student would tend to call more frequently if the price of telephone calls was lower.

Recall that the demand curve was derived by holding income, tastes, and other prices constant. If that condition (*ceteris paribus*) were relaxed, we would have to derive an entirely new relationship between price and quantity. For example, consider increasing the income available to our graduate student. If she had higher income, it seems likely that she would call home more frequently at whatever the price turns out to be.

Figure 2.3 presents such a change for our hypothetical graduate student. At a price of 50 cents, she would increase the frequency of calling from twenty-three to twenty-seven calls per month. At $3.50 per call, she would increase the frequency of calling from seven calls per month to eighteen calls per month. Similarly, at a price of $10 per call quantity demanded would increase to seven calls per month. Notice that if calls are free, the amount of income she has doesn't matter; at a zero price, demand does not increase.

Such a change is often referred to as a *shift in the demand curve.* It is important to remember that what really is transpiring is that the conditions that were in place at the time the original demand curve was derived have now changed. We have derived a new relationship between price and quantity.

The shape of any demand curve and its position depends critically on the income of the household and its preferences. Nothing can be presumed about the character of those preferences. Some people have a strong taste for talking on the telephone, and others may have very high incomes. Some may have a strong preference for the written word. Still others may have very low incomes, which may make even a price of $1 seem high. The diversity of individual demands is infinite.

Market demand is simply the sum of all the amounts demanded by all the households shopping in the market in question. Figure 2.4 shows

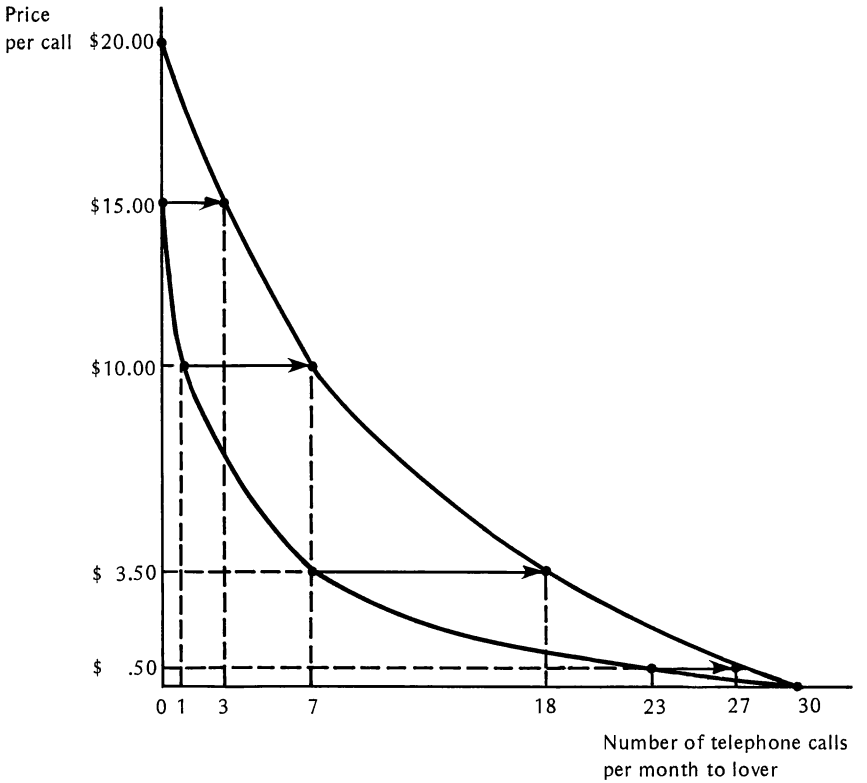


Figure 2.3. *Shift of a demand curve following a rise in income.*

the derivation of a market demand curve from individual demand curves. The derivation is for only three people: Most markets, of course, have thousands of demanders.

At a price of $3.50 both *A* and *C* would purchase 4 pounds of coffee per month, while *B* would buy none, presumably switching to tea. Market demand would thus be a total of 4 plus 4 or 8 pounds. Similarly at a price of $1.50, *A* would purchase 8 pounds, *B*, 3 pounds, and *C*, 9 pounds per month; thus at a price of $1.50 market demand would be 8 plus 3 plus 9 or 20 pounds of coffee per month. Effective demand in the marketplace is nothing more than the sum of all the purchases of all the individuals shopping in the market. A market demand curve shows the amount of a product that will be sold at every price if households could buy all they wanted at that price.

As you can see from the diagram, market demand is the *horizontal sum* of all the individual demand curves. The market demand curve

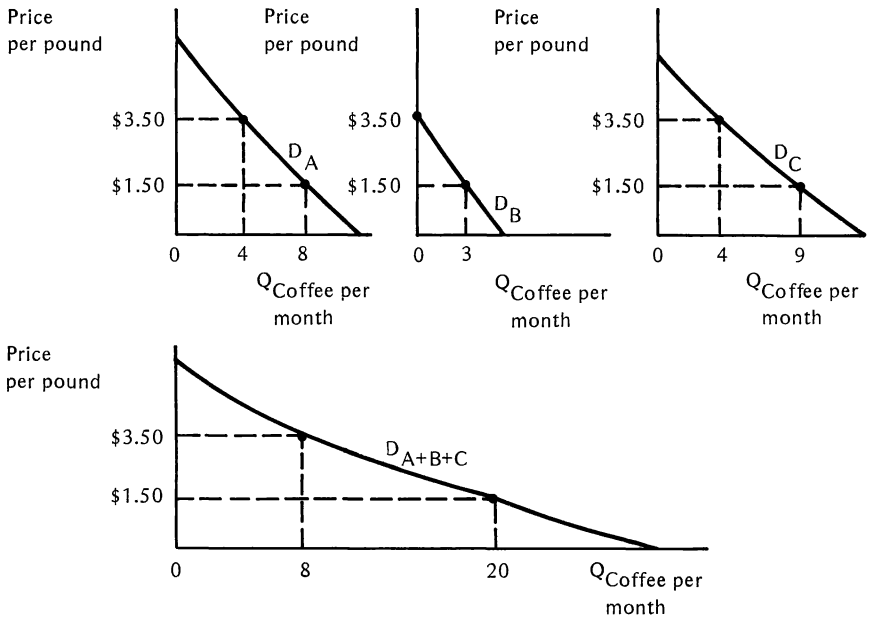


Figure 2.4. *Deriving market demand from individual demand curves.*

thus takes its shape and position from the shapes, positions, and number of individual demand curves. If more people decide to shop in this market, more demand curves must be added, and the market demand curve will move to the right. Market demand curves may shift as a result of preference changes, income changes, or in the number of demanders.

Business Firms and Supply in Output Markets

A second major portion of microeconomic theory attempts to analyze the behavior of business firms. Business firms are economic agents that engage in production, and it is assumed that they do so for profit. Profit is, in very simple terms, the difference between revenues and costs. That is, successful firms are able to sell their product for more than it costs to produce it.

The amount that an individual firm will produce or *supply* of a product depends on a number of factors:

1. The price at which the product can be sold;

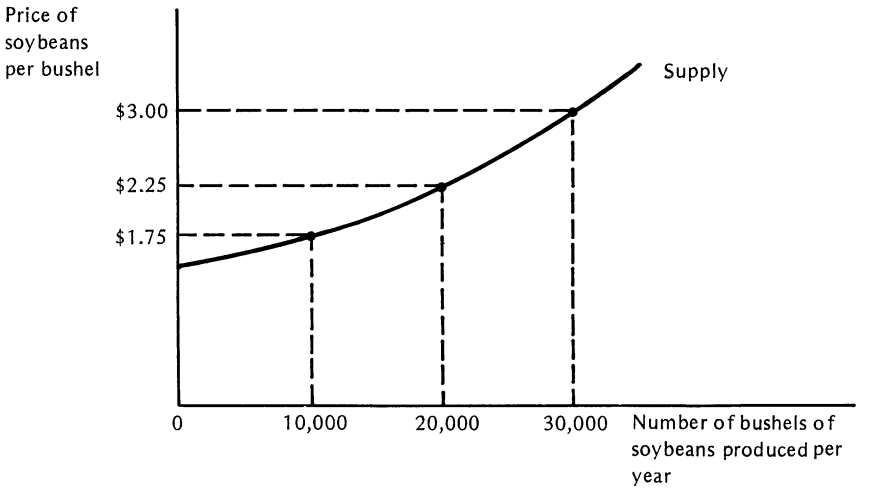


Figure 2.5. *Individual supply curve.*

2. The cost of producing the product;
   a. The prices of inputs required to produce it;
   b. The quantity of each input required to produce output (the technology of production).

In their decision to supply a product, firms must analyze a good deal of data. Most products can be produced using a variety of different techniques and combinations of inputs. Digging a hole for a foundation of a building can be done with large numbers of men and women with shovels or by a small number of men and women with steam-shovels. The final cost of any product thus depends on the techniques available as well as the cost of inputs.

The following chapters analyze the behavior of firms in more detail. Here we focus on the relationship between output price and quantity supplied *ceteris paribus*. A *supply curve* is a graph of the relationship between the quantity supplied of an individual product by a single firm and the price of that product. It is reasonable to assume that at higher prices firms are likely to produce and thus supply more of the product.

Figure 2.5 presents a hypothetical supply curve for an individual farmer. The relationship suggests that if price were to rise from $1.75 per bushel to $3 per bushel, the quantity of soybeans supplied *per period* in bushels would triple from 10,000 to 30,000. The relationship is derived holding the state of technology and input prices at constant levels.

The shape and position of an individual supply schedule depends on the cost of producing the product. Cost, in turn, depends on the

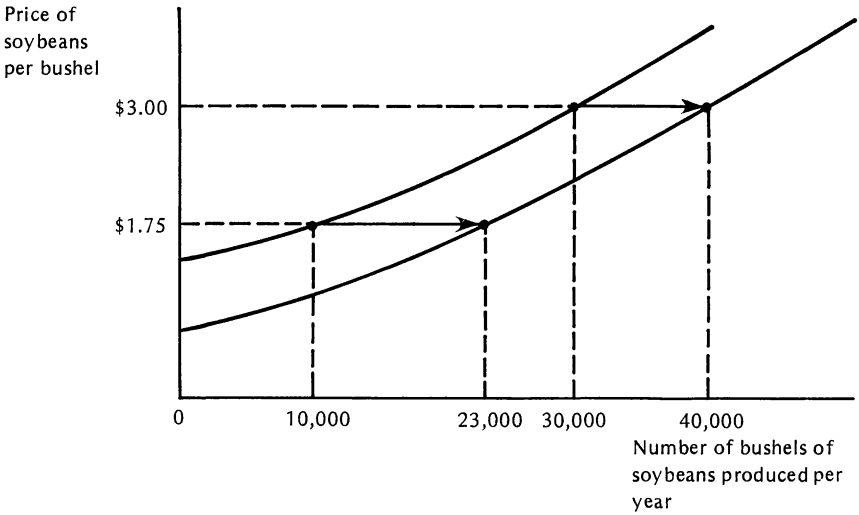


Figure 2.6. *Shift of a supply curve following a fall in cost of production.*

price of inputs and the techniques available for producing the product. If technology were to change or input prices were to adjust, cost of production would change and the supply curve of a firm might shift.

Let us suppose for an example that a major breakthrough is made in the production of soybeans—such as an advance in genetic engineering or the availability of a new fertilizer. Such a change might induce individual farmers to supply more soybeans at any given market price. Figure 2.6 presents a diagram of this kind of a change.

At a price of $3, farmers would have produced 30,000 bushels; with the lower cost of production that number would increase to 40,000. Similarly, at $1.75 per bushel farmers would have produced 10,000 bushels, at the lower cost of production that output would rise to 23,000 bushels.

Market supply, as with market demand, is simply the sum of all that is supplied by all producers in an individual market. Figure 2.7 presents the derivation of a market supply curve from the supply curves of individual firms. If there were only three firms, market supply would be the sum of the amounts produced by each. At a price of $3, farm *A* would supply 30,000 bushels of soybeans, farm *B* would supply 10,000 bushels, and farm *C* would supply 25,000 bushels. As a result, the total amount supplied in the market would be 30,000 plus 10,000 plus 25,000 or 65,000 bushels. At a price of $1.75, the total amount supplied would be 25,000 bushels. Thus, the market supply schedule is nothing more or less than the simple horizontal addition of all the individual firms' supply schedules.

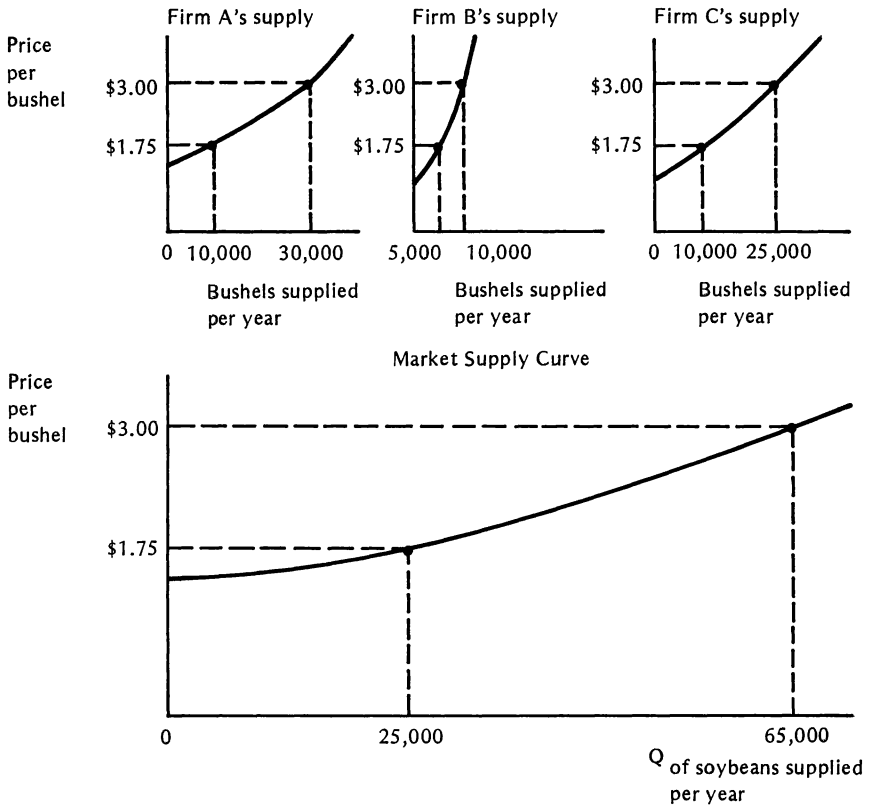


Figure 2.7. *Deriving market supply from individual firm supply curves.*

The Nature of Market Equilibrium

In our discussion of demand and supply, we have very carefully separated the factors that influence household decisions about how much to demand and the factors that influence firm decisions about how much to produce. The operation of the market, however, depends on the interaction of suppliers and demanders.

At any moment, one of three conditions will prevail in any market: (1) The quantity supplied will equal the quantity demanded; (2) the quantity demanded will exceed the quantity supplied (a situation called a *shortage*); or (3) the quantity supplied exceeds the quantity demanded (a situation called a *surplus*).

When quantity demanded exceeds quantity supplied, it is reasonable to expect that price will rise; economic theory assumes that some orderly mechanism will allow those willing to pay more to bid the prod-

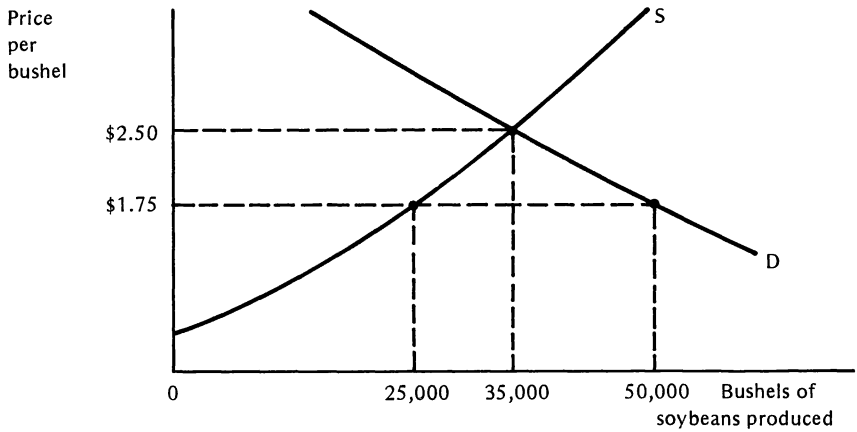


Figure 2.8. *A shortage.*

uct away from those willing to pay less. When the opposite is true and producers discover that more is being produced than is being sold, there is a tendency for prices to fall; once again, economic theory assumes that prices will fall in response to a surplus. Finally, when quantity supplied and quantity demanded are equal, it is reasonable to anticipate that price will remain stable, a situation referred to as *equilibrium.*

Figure 2.8 illustrates a shortage. At a price of $1.75 per bushel producers are supplying 25,000 bushels of soybeans per year, while demanders would purchase 50,000 at the price. If, as theory postulates, price began to rise above $1.75, two things would happen. First, the higher price would induce producers to produce more soybeans. Second, in the face of higher prices demanders would shift toward other products or simply cut down on their use of soybeans. Thus, quantity supplied would increase, and quantity demanded would decline until an equilibrium was reached at a price of $2.50. At that price, 35,000 bushels of soybeans would be exchanged on the market.

Figure 2.9 illustrates a surplus. At a price of $3 per bushel the quantity supplied is 65,000 bushels. At the same price consumers would purchase only 20,000 bushels. As a consequence, theory postulates that price tends to fall as producers find inventories building up. If price beings to fall, two things happen: (1) Producers cut back on the quantity of soybeans being produced per year; and (2) buyers increase their purchases because of the lower price. As price falls, quantity demanded increases, and quantity supplied decreases, until, again, equilibrium is reached at a price of $2.50 and a quantity of 35,000 bushels.

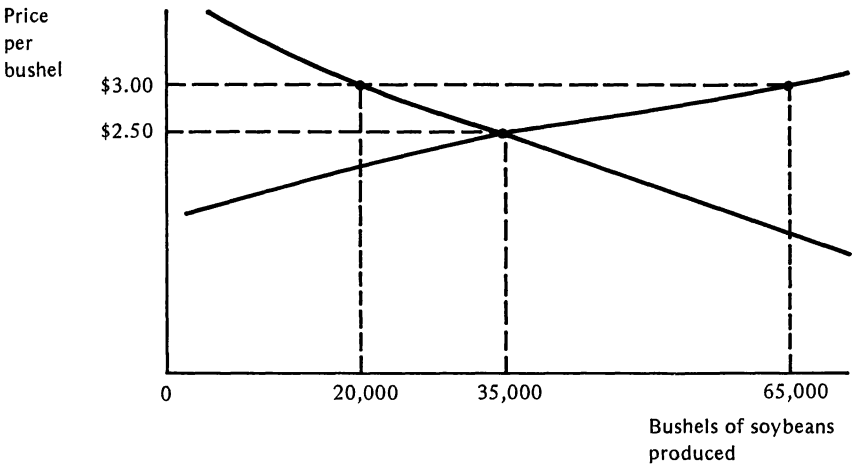


Figure 2.9. *A surplus.*

A Change in Equilibrium Resulting from a Shift of Demand or Supply

In the spring of 1985 those concerned with the coffee market were paying a great deal of attention to the weather in Brazil. Brazil is one of the major producers of coffee beans in the world. A cold snap in Brazil can reduce the coffee harvest by enough to affect the world price of coffee beans. In the mid-1970s a major freeze in South America had such an effect and drove the price of coffee in grocery stores in the United States from around $1.50 to $4.50 per pound. Figure 2.10 illustrates a shift in supply resulting initially in a shortage. At the initial price of $1.50 the quantity supplied declines sharply, and quantity demanded (*Qd*) is greater than the quantity supplied (*Qs*). Assuming that the market is unregulated, the price will begin to rise. As it rises, recall that two things occur: (1) Quantity demanded declines as people shift to substitute products or decide that coffee is too expensive; (2) the quantity supplied begins to increase but within the limits imposed by the damage from the freeze—that is, supply increases along the new supply curve, which lies to the left of the old.

The mechanism by which price rises in response to a shortage, or falls in response to a surplus, differs dramatically from market to market. In some markets, such as an auction, prices adjust instantaneously to eliminate shortages and surpluses and to clear the market. The federal funds market is one in which banks loan cash to other banks for very short periods of time. Those requiring extra funds and those with

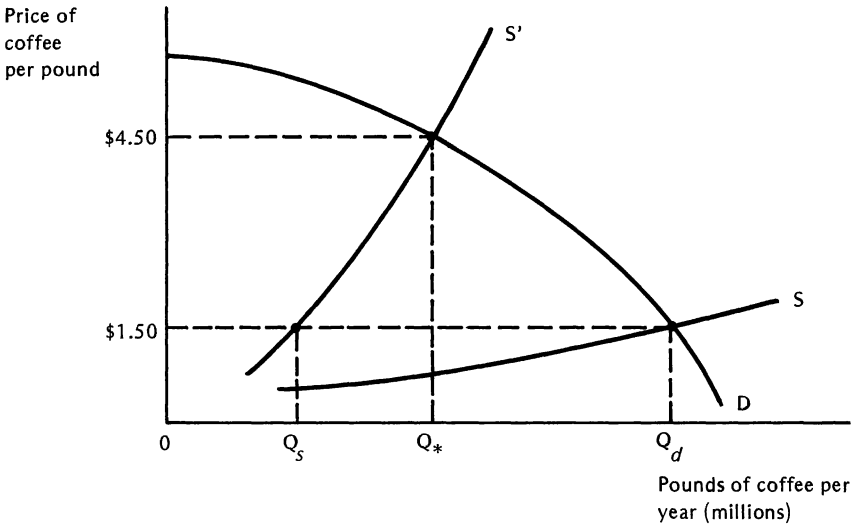


Figure 2.10. *A shift of supply and subsequent price adjustment.*

surplus funds to lend contact a central desk in New York, and the interest rate (the price that borrowers pay to lenders) adjusts literally on a minute-by-minute or hour-by-hour basis. In the housing market, information is much more difficult to come by, and adjustment takes more time. Chapter 3 discusses the nature of this rationing system in more detail.

Important Things to Remember About the Mechanics of Supply and Demand

1. Demand is determined by household preferences that are constrained by prices and by household income and wealth.
2. Supply is determined by selling price and cost of production. Cost of production is determined by input prices and available technologies of production.
3. We are asking how much the household would buy if it could buy *all it wanted* at the given price. On the supply side, the question is how much would be produced if a firm could sell *all it wanted* at a given price.
4. Remember it is *price per unit.*

5. Quantity demanded is *per time period* (that is, per day, per month, or per year).
6. Be careful to distinguish between movements along supply and demand curves and shifts of those curves. For example, when the price of a good changes, the quantity demanded (or supplied) of that good will change; that is a movement along the curve. Since the curves were derived *ceteris paribus*, when any of the other determining factors change, the curve may shift or change position. It is as if we derived another curve. For example, a change in income may shift the demand curve to the left or right; a change in cost of production may shift a supply curve to the left or right.

Review Questions and Exercises

1. Think of some commodity that you like (try to avoid "lumpy" things like cars or houses, and pick something like coffee, movies, trips to a frequently visited place, records, long distance phone calls, and so forth.
   a. Roughly sketch your demand curve for that commodity. Does it hit the price axis? Where? How much would you buy at a zero price?
   b. Are there substitutes for this commodity? How does the availability of substitutes affect the shape of your curve?
   c. How would your demand curve change in response to the change in the price of some substitute?
   d. How would it change if you won the lottery and were to receive $1,000 per week for life?
2. The U.S. government administers two programs that affect the market for cigarettes: Media campaigns and labeling requirements are aimed at making the public aware of the dangers of cigarettes. At the same time, the Department of Agriculture maintains a program of price supports for *tobacco*. Under this program, the supported price is above the market equilibrium price, and the government limits the amount of land that can be devoted to tobacco production. Are these two programs at odds with respect to the goal of reducing cigarette consumption? Explain carefully. As a part of your answer illustrate graphically the effects of both policies on the market for cigarettes.
3. There was a good deal of evidence in the 1960s that the urban

housing market was in fact a "dual" housing market divided along racial lines—that is, there were completely separate supply and demand forces at work in black and white submarkets. Blacks could buy housing only in certain areas, and the supply of housing in these areas was relatively fixed. In addition, the empirical evidence indicated that the price of housing in the black submarket was higher than in the white submarket: Houses in the black area were selling for more than identical houses in the white area. Urban economists attributed this differential to *demand pressure,* since demand trends showed that blacks had been moving from the rural South to the urban North in large numbers. A recent economic study of discrimination in the rural South found the same sort of dual housing market. Since rural black populations have been declining both relatively and absolutely, the research team anticipated finding a price differential favoring blacks (that is, that blacks pay a lower price for housing). Instead, they found that the price of housing in black neighborhoods was in fact *higher* than the price of identical housing in white neighborhoods despite the decline in black population. Give at least two possible explanations for this seeming paradox. Use simple supply and demand graphs in your explanation.

4. Assume that the Boston Red Sox baseball team charges $5 per ticket for all seats to all regular season games. Assume also that the capacity of their stadium (Fenway Park) is 35,000. In August the Red Sox played games against the New York Yankees (a great rival) and against the Cleveland Indians (a team in last place) on consecutive Sundays. All tickets to the Yankees game were sold out a month in advance, and many people who wanted tickets could not get them. At the same time there were 15,000 vacant seats at the Cleveland game.
   a. Draw supply and demand curves for tickets to the two games (two graphs) as you might imagine them.
   b. Is there a pricing policy that would fill the stadium for the Cleveland game? Would such a policy bring the Red Sox more or less revenue?
   c. Since the price rationing system is not used to ration tickets to the Yankee game, how might tickets be rationed?

3

Supply and Demand: The Nature of the System, Tax Applications, and the Concept of Elasticity

As was described in Chapter 1, every society has a system of institutions that determine what gets produced, how it gets produced, and for whom. While in many societies decisions about these economic outcomes are made centrally through planning agencies or government directive, in every society many decisions are made in a decentralized way through the operation of the market system. Markets exist in all societies.

While Chapter 2 described the bare bones of market operations, this chapter reflects briefly on the nature of the outcome. It then discusses an important tax policy application and finally the important concept of *elasticity*.

The price system works as a mechanism to allocate scarce resources among alternative uses. Consider, first, the simple process by which the price system eliminates a shortage. Figure 3.1 illustrates the situation that existed at the time of the coffee freeze in the mid-1970s. At the initial price of $1.50, there was a significant excess demand. There was, quite simply, not enough coffee to satisfy all that was demanded at that price.

The shortage was quickly eliminated as price rose. Look carefully at what occurred as price increased from $1.50 to $4.50. While there was some increase in coffee output as the higher price attracted producers, the most significant factor was the reduction in demand. As price rose, the number of pounds that households wanted to pur-

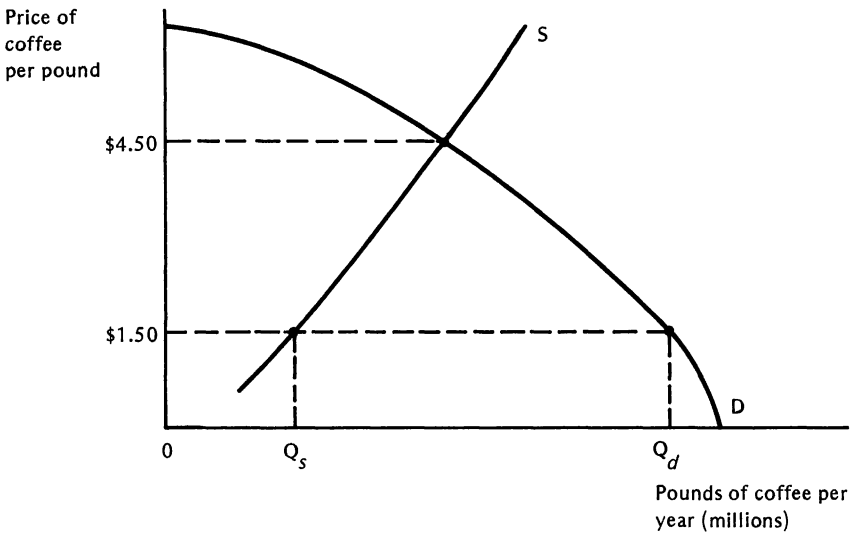


Figure 3.1. *Price rationing in the coffee market.*

chase declined. Some households simply found that coffee at the new higher price was too expensive. Others were able to switch to substitutes, such as tea or hot chocolate. The most important thing to note about the decline in demand is that those who were willing to pay the higher price obtained the coffee. Willingness to pay, of course, is constrained by ability. One's ability to purchase a commodity at a higher price depends on income and wealth, which is why some people object to the price rationing system: It allocates goods and services in accordance with income and wealth.

There is some price that will clear any market you can think of. Consider the market for a famous painting, such as the Mona Lisa. Figure 3.2 illustrates the operation of such a market. At a low price, there would be an enormous demand. Supply is simply one unit. The price will be bid up until there is only one remaining demander. The demander who gets the Mona Lisa is the one who is willing and able to pay the most. Presumably, that price will be very high. If the product is in strictly scarce supply, such as a single painting, its price is said to be *demand determined*—that is, it is determined exclusively by the amount that the highest bidder, or highest bidders, are willing to pay.

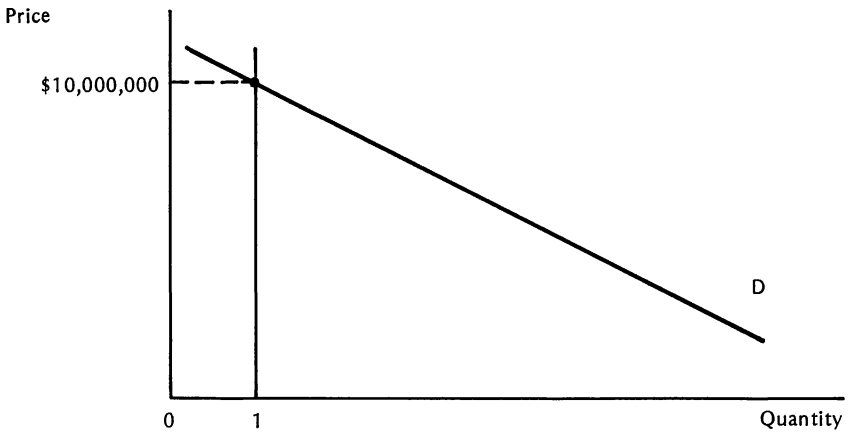


Figure 3.2. *Market for a single copy of a rare painting.*

Constraints on the Market and Alternative Rationing Mechanisms

Governments as well as private agencies may decide not to use the market system for rationing an item in short supply. By short supply, of course, we mean quantity demanded in excess of quantity supplied at the current price. If price is not allowed to rise to equilibrium, another mechanism must be found to divide up the shortage.

In 1973 and 1974 the Organization of Petroleum Exporting Countries (OPEC) succeeded in raising the price of crude petroleum on world markets by over 400 percent. In addition, they imposed an embargo on shipments of crude oil to the United States, and there ensued a dramatic shortage of gasoline at the pump. Had the market system been allowed to operate, the price of refined gasoline would have increased very dramatically. Prices would have increased until the point that quantity supplied was equal to the quantity demanded. Those who were willing and able to pay a very high price would have been the ones to get the gasoline. In this instance, the government decided that rationing gasoline only to those who were able to pay the most—that is, the rich—was unfair. As a result, the Congress imposed a price ceiling, or maximum price, on gasoline at the pump—a price ceiling was intended to keep gasoline "affordable."

It also, however, perpetuated the shortage. At the restricted price, quantity demanded remained greater than quantity supplied. But somehow that gasoline had to be divided among all potential demanders. The most common of all non-price rationing systems is queueing (which

simply means standing in line). During 1974 people often waited for hours in very long lines at gas stations, starting as early as 5:00 A.M. Under this system, gasoline went to those who were willing to pay the most, but the sacrifice was measured in hours and in annoyance.

A second non-price rationing device is the device of the *favored customer*. Many gas station owners decided not to sell gasoline to the general public at all but to reserve their scarce supplies for friends and favored customers. Not surprisingly, many customers tried to become favored by offering side payments to owners of gasoline stations.

Another method of dividing up a shortage is to use ration tickets. It was suggested both in 1974 and 1979 that each family be given the right to purchase a certain number of gallons of gasoline each month. The entitlement would be distributed to families in the form of ration tickets or coupons. Such a system was employed during the 1940s when price ceilings on many items had been imposed.

It is important to note that if ration coupons were used but there was no prohibition against trading those coupons, the result would be almost identical to a system of price rationing. That is, those who were willing and able to pay the most would simply buy up the coupons and use them to purchase gasoline at the restricted price. Even when the trading of coupons is declared illegal, it is difficult for the authorities to stop the development of *black markets*.

Important sporting events such as the World Cup or the Super Bowl take place in stadiums with finite capacities. In most cases, there are large numbers of fans who would like to attend those events. One way of rationing tickets, of course, is for the sponsors to sell the tickets at their market values. To an event like the finals of the World Cup, the price that clears the market could be astronomical. In 1975, for example, the Boston Red Sox baseball team was involved in the World Series with the Cincinatti Reds. Estimates done at the time suggested that the Red Sox organization could have sold all tickets to the World Series for over $50 per ticket. Tickets, however, were actually sold for a price far below equilibrium, and there was a shortage in the classic sense: The quantity demanded exceeded the quantity supplied at the price being charged.

The Red Sox organization, of course, had to decide to whom the tickets would be sold, and several devices were employed. First, they devised a lottery. They would take all letters postmarked after midnight on a specified date, put them in a large bin, and select the lucky recipients randomly. The justification for using the lottery was that it would be unfair to let only rich people go to the games. Using the price system would allocate the tickets not to the loyal fans who had come to regular season games at low prices but rather only to the rich.

In fact, tickets to the Red Sox World Series were divided up in a

variety of ways: Some went to season ticket holders; some, to winners of the lottery, and some, to state legislators, sponsors, and friends of the Red Sox organization. Many who won the lottery sold their tickets to those who were willing to pay high prices: a very extensive black market developed. At one point an entire page of ads in the local newspaper offered World Series tickets at very high prices.

The point of this discussion is to describe the nature of the price rationing system and to suggest some alternatives. No matter how well-intentioned private organizations and governments are, however, it is very difficult to prevent the price system from operating. Every time an alternative is proposed, the price system seems to sneak in the back door.

Applications of Supply and Demand to Tax Policy: The Incidence of an Excise Tax in the Short Run

It doesn't take a great deal of analysis to show that the person or institution that initially pays a tax does not necessarily bear the burden of that tax. The classic example is the corporate income tax. The corporate income tax is, of course, paid by corporations. But corporations as institutions have no tax-paying capacity. Corporations are extensions of individuals, and the question really is which individuals ultimately end up paying the tax. There are several candidates. Taxes imposed on corporations may be ultimately paid by owners of corporations. They may also be paid by consumers of corporate products or by employees. Makers of tax policy must know who will ultimately pay the taxes that are imposed.

The ultimate distribution of the burden of any tax depends on the ways in which markets react or adjust to the imposition of those taxes. This section provides a simple example of tax shifting that takes place as a result of supply and demand adjustment in a very simple model.

Figure 3.3 shows equilibrium in a market with no tax. Quantity supplied by all firms in a market depends on the price received by firms for the product. Demand by all households in a market depends on the price that households must pay to obtain the product. When there is no tax on the item in question, the amount that consumers pay is exactly the same as the amount that producers receive, and the market will clear as that price adjusts. The presence of a tax, however, whether imposed on producers or consumers directly, means that the amount that consumers pay and the amount that producers receive are different. If a tax is imposed directly on consumers, then they must pay it in addition to the net price of the product, and thus the total

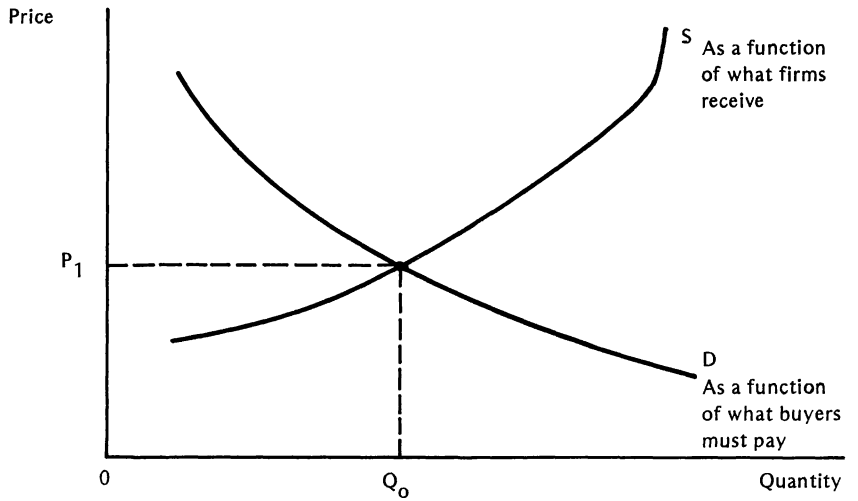


Figure 3.3. *Equilibrium in a market with no tax.*

price includes the tax. If the price is imposed on producers, producers must take the tax out of their receipts and the net price that they receive is lower than the one that is paid by consumers.

The presence of a tax means that we can draw the supply curve in two ways. In Figure 3.4 we begin with supply as a function of what producers receive. It is a standard supply curve, which reflects the behavior of firms in response to changing prices directly. In the diagram, if producers receive \$5, they will supply 11,000 units; similarly, if they receive \$3 per unit, they will supply 5,000 units per period. The imposition of a tax means that we can draw the same supply curve as a function of what consumers pay. In order for producers to receive \$5, consumers must pay \$6. If consumers pay \$6, producers will receive \$5 and thus supply 11,000 units. For producers to receive \$3 per unit, households must pay \$4. If households pay \$4, producers will receive \$3 per unit and thus supply 5,000 units per period.

Figure 3.5 shows the adjustment that would take place in response to the imposition of an excise tax of \$1 that must be paid by producers directly. Initially, the market is in equilibrium with a price of P_1 and the market clearing at a quantity of Q_1. Imposing a tax of T on producers means that producers will now receive only $(P_1 - T)$. If they receive only that amount, the quantity that they would supply is reduced to Q_3. At that point, quantity demanded remains at Q_1, but quantity supplied has been reduced to Q_3. Thus, there is a shortage in the market, with quantity demanded in excess of quantity supplied.

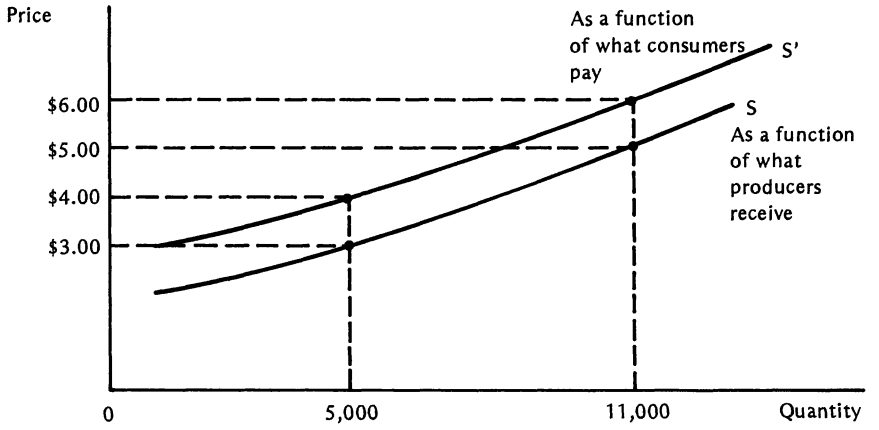


Figure 3.4. *Two supply curves when a tax is levied on producers.*

In response to that shortage, the market price begins to rise. It will rise all the way to P_2. Notice that at P_2 supply and demand are again equal; the quantity demanded is Q_2. At P producers will receive $P_3 = (P_2 - T)$. At that price producers will also produce a quantity Q_2. Thus, at a market price of P_2, the market will clear.

Notice that the imposition of a tax has resulted in an increase in price but that price to consumers has not risen by the full amount of the tax. Equilibrium price rises from P_1 to P_2. The amount received by producers for each unit produced, however, falls from P_1 to P_3. The increase in price to consumers plus the decrease in return per unit to producers equals the total amount of tax per unit of product. Thus, the "burden" of the tax is shared by producers and consumers. Part of the tax is paid by consumers in the form of higher prices; part of the tax is paid by producers in the form of lower receipts.

Notice in Figure 3.5 that the *total* tax paid is equal to $T \times Q_2$ or simply the tax per unit times the number of units sold, which is equal to the area of rectangle $P_2P_3\,BA$. The consumers' share of the tax is equal to the price rise per unit $(P_2 - P_1)$ times the number of units sold, Q_2. Thus, the consumer share is equal to the area of rectangle $P_2P_1\,CA$. The producer's share of the tax is equal to the decline in the return to the producers *per unit* $(P_1 - P_3)$ times the number of units sold, Q_2. Thus, the producers' share is equal to the area of rectangle $P_1P_3\,BC$.

The relative size of the two shares of burden depends on the shape of the demand curve and the shape of the supply curve. Figure 3.6a illustrates that when demand is relatively unresponsive to price (inelastic), consumers bear the lion's share of the burden. Figure 3.6b

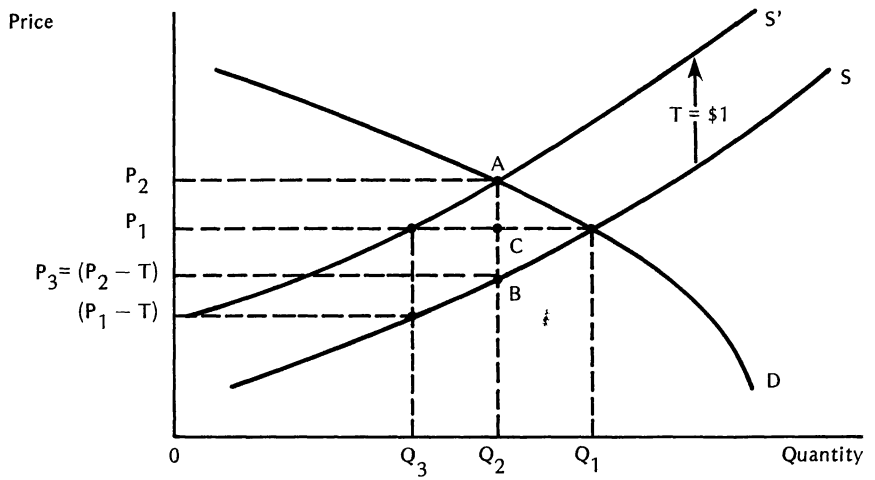


Figure 3.5. *Shifting of an excise tax in the short run.*

illustrates that when the higher price results in a very sharp drop-off in quantity demanded (elastic demand), producers bear the lion's share of the burden. The concept of elasticity will be explored in more detail in the next section of the chapter.

Elasticity

The last two chapters have discussed in some detail the basic principles of supply and demand. Those principles are theories about the ways in which households and firms are likely to behave. For example, when prices rise, households are likely to purchase less and firms are likely to supply more of a product. When cost of production falls, supply is likely to shift outwards.

As the tax example above illustrates, the size or magnitude of these reactions can be very important. If the entire world could have switched to natural gas when OPEC raised the price of crude oil, raising the price of petroleum would not have been such a wise move. It would probably not be wise for the organization of banana exporting countries to try it.

The U.S. government spent hundreds of millions of dollars on research to determine whether the current welfare system actually deterred people from entering the labor force and working. In economic language, they wanted to know whether an increase in the *net wage* (price of labor) would increase the supply of labor and if so by how much.

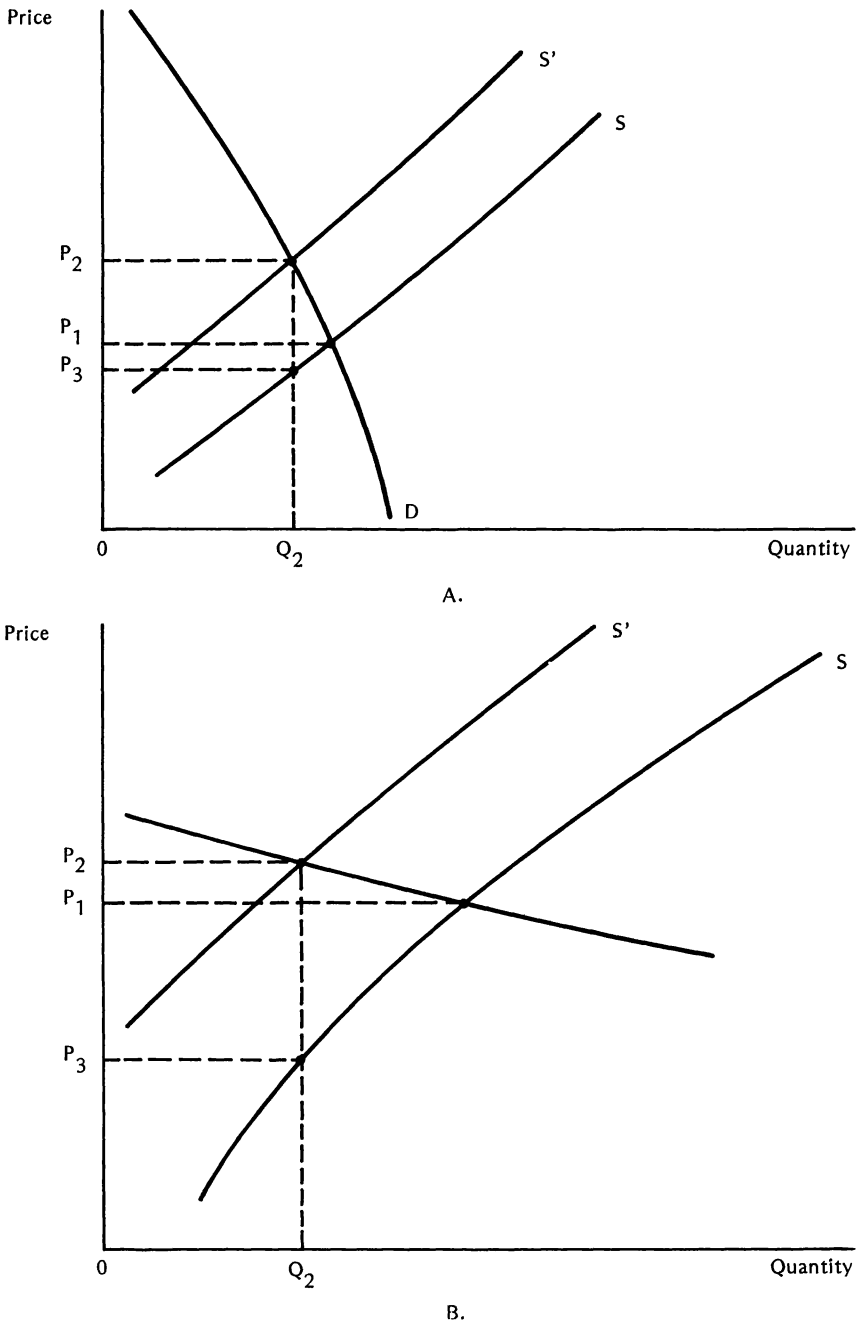


Figure 3.6 (a,b). *An excise with elastic and inelastic demand.*

Similarly, millions were spent to determine how the supply of privately produced rental housing was likely to respond to a shift in housing demand induced by a subsidy paid to households. There are thousands of other examples.

In the last twenty years, the discipline of economics has changed dramatically. A substantial portion of the research being done in economics today involves the collection and analysis of quantitative data designed to "measure" behavior. To measure the size of a response, economists generally employ the concept of *elasticity*. It is a very general concept. If some measure or *variable* causes a response in another measure or variable, elasticity is a measure of the size of the response. We define the elasticity of B with respect to changes in A as the percentage of change in B divided by the percentage of change in A:

$$\frac{\%\ \Delta\ B}{\%\ \Delta\ A}$$

The price elasticity of demand is defined as the percentage of change in quantity divided by the percentage of change in price:

$$\frac{\%\ \Delta\ Q_d}{\%\ \Delta\ P}$$

The diagrams in Figure 3.7 present two identical supply shifts. In both cases a shortage is created by the old price, and price begins to rise. As it does, demand responds, but the demand decline is greater and more rapid in B than in A. To calculate elasticity in A:

$$\%\Delta\ Q = \frac{Q_2 - Q_1}{Q_1} \cdot 100 = \frac{90 - 100}{100} \cdot 100 = -10\%$$

$$\%\Delta\ P = \frac{P_2 - P_1}{P_1} \cdot 100 = \frac{9 - 5}{5} \cdot 100 = 80\%$$

$$\text{Elasticity} = {}^{-10}/_{80} = -{}^{1}/_{8}$$

To calculate elasticity in B:

$$\%\Delta\ Q = \frac{Q_2 + Q_1}{Q_1} \cdot 100 = \frac{50 - 100}{100} \cdot 100 = -50\%$$

$$\%\Delta\ P = \frac{P_2 - P_1}{P_1} \cdot 100 = \frac{6 - 5}{5} \cdot 100 = 20\%$$

$$\text{Elasticity} = {}^{-50}/_{20} = -2{}^{1}/_{2}$$

When the absolute value of elasticity is greater than 1, demand is said to be elastic (responsive). When the absolute value of elasticity

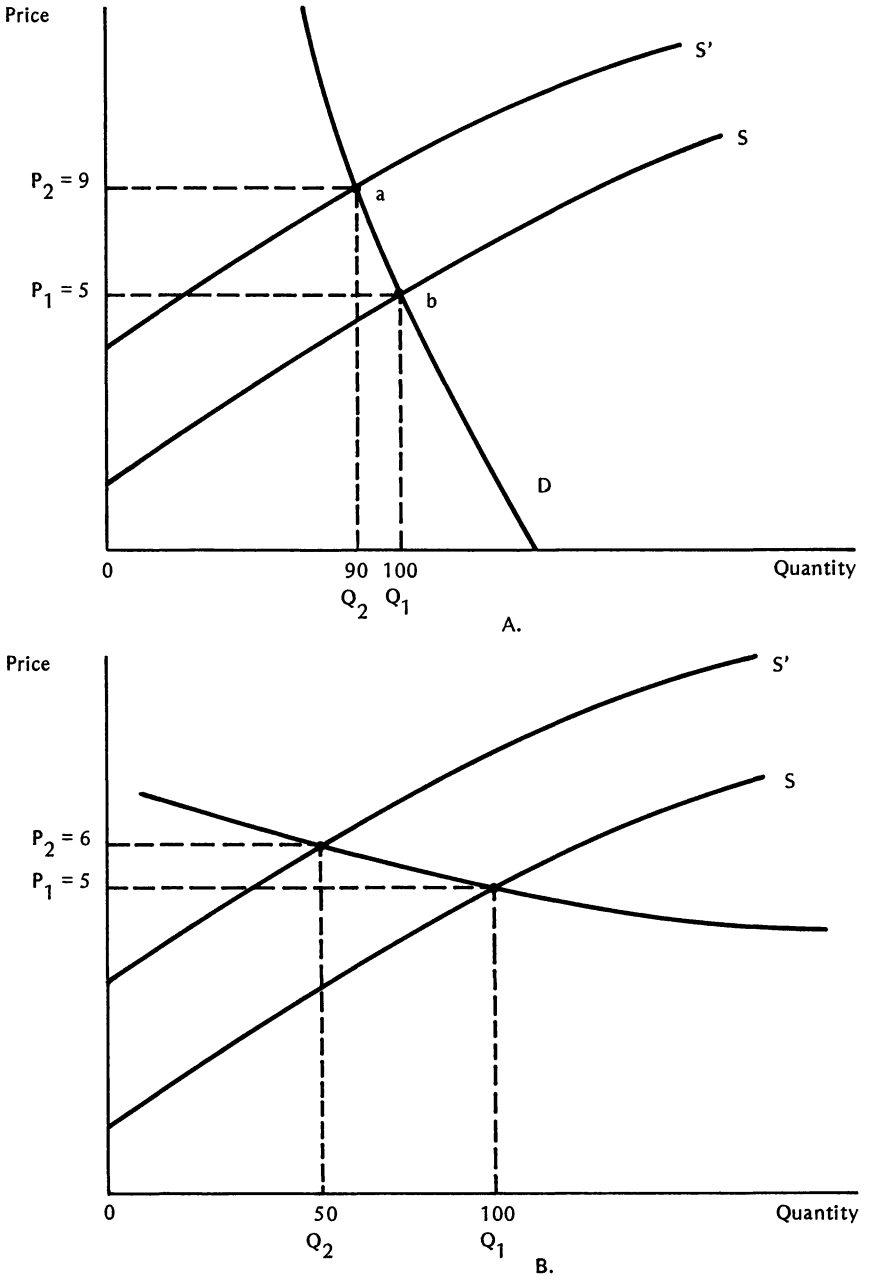


Figure 3.7 (a,b). *Two different responses from an identical shift of supply.*

is less than 1, demand is said to be inelastic (relatively unresponsive).

Notice that as defined, elasticity may be positive or negative. Price elasticity is negative, since an increase in price (a positive change) leads to a decrease in quantity demanded (a negative change).

If a good is a normal or luxury good, an increase in income will lead to additional demand; for such goods the elasticity of demand with respect to income is positive.

Elasticity and Total Expenditure or Revenue

In any market $P \times Q$ is total expenditure on the good as well as total revenue to producers. The total amount spent by a household on a good is $P \times Q$. Total expenditures may rise or fall when prices change. When prices *rise*, quantity demanded *falls*:

$$P^x \uparrow \rightarrow Q_D^x \downarrow$$

If the decline in demand is *smaller* than the increase in price, total expenditure or revenue will rise when prices rise. This occurs if demand is inelastic:

$$\uparrow P \cdot Q \downarrow = TR \uparrow$$

The price rise outweighs the quantity decline, and P x Q rises.

If the decline in demand following a price increase is larger than the increase in price, total expenditure/revenue will *fall*. This occurs if demand is elastic:

$$\uparrow P \cdot Q \downarrow = TR \downarrow$$

The price rise is swamped by the quantity decline, and *P x Q falls*. The reverse is true for a price cut. For a summary see Table 3.1.

Review Questions and Exercises

1. For the past ten years the price of natural gas sold in interstate commerce has been regulated by the federal government. Currently the regulated price is thought to be substantially below equilibrium. For several years the Congress debated those regulations.

Table 3.1
Changes in Total Expenditures
Revenues in Response to Price Change

	Absolute Value of E_D			
	Elastic $E_D > 1$	*Unit Elastic* $E_D = 1$	*Inelastic* $1 > E_D > 0$	*Perfectly Inelastic* $E_D = 0$
Effect of fall in *P* on (*P* × *Q*)	$\frac{\triangle Q}{Q} > \frac{\triangle P}{P}$	$\frac{\triangle Q}{Q} = \frac{\triangle P}{P}$	$\frac{\triangle Q}{Q} < \frac{\triangle P}{P}$	$\triangle Q = 0$
	Increase	Constant	Decrease	Fall in (P × Q) is proportional to fall in P
Effect of rise in *P* on (*P* × *Q*)	Decrease	Constant	Increase	Rise in (P × Q) is proportional to rise in P

NOTE: To the business firm, P × Q = TR *(Total Revenue).*

a. If prices have been held at their current levels for a long time, how do we know they are not at an equilibrium? What evidence of a disequilibrium would you look for?

b. Illustrate the current situation with supply and demand curves.

c. Those who argue for deregulation claim that deregulation will increase the supply of natural gas. Those in favor of maintaining the regulations argue that the supply of natural gas is very *inelastic.* If supply is very inelastic and thus will not respond to high prices, what would happen if price was deregulated?

2. Taxi cab fares in most cities are regulated. Several years ago Boston taxi cab drivers obtained permission to raise their fares 10 percent. They anticipated that revenues would increase by about 10 percent as a result. They were, however, disappointed. When the commissioner granted the 10 percent increase, revenues increased by only about 5 percent. What can you infer about the elasticity of demand for taxi cab rides? What were taxi cab drivers assuming about the elasticity of demand?
3. The freeze that destroyed a good portion of the South American coffee crop in the mid-1970s increased the price of tea. Explain using supply and demand diagrams.
4. Several members of the law school faculty were standing in a long line at the student cafeteria. One was heard to remark that she wished the cafeteria would increase prices. Can you explain why?
5. Review the section on the incidence of an excise tax in the short run. Notice that the imposition of an excise tax is borne in part by consumers of the product in the form of higher prices and in part by producers of the product. Suppose that the lower price received by producers led to lower profits and as a consequence a number of the firms in the industry eventually went out of business. Use diagrams to explain what would happen to the burden of the tax in the long run.

4

Household Behavior and Consumer Choice

The decision of a household to purchase or demand some amount of a good is one of a series of household decisions that are closely related. This chapter briefly explores the economics of household choice.

Chapters 4 through 7 together build a *model* or *theory* about the operation of a simple competitive market economy. The term *competitive* means that we assume that both households and firms are *price-takers* in input and output markets. No individual household and no individual firm has control over market price; price is determined by the interaction of many firms supplying and many households demanding. Clearly, this assumption is realistic for some sectors of the economy, such as agriculture, but not for others. While an individual farmer has no power over the price of wheat and individual consumers have no power over the price of bread, IBM has something to say about the price of personal computers. Alternative noncompetitive market structures will be explored later in the book.

The Idea of Constrained Choice

In the chapter on supply and demand, it was argued that several factors influence the quantity of a good that a given household would demand:

1. The income available to the household;
2. The amount of accumulated wealth of the household;
3. The price of the product in question;

4. The prices of other products available to the household;
5. The tastes or preferences of the household.

If we begin by taking wealth, income, and prices as given, items 1 through 4 simply define the set of choices that are available to a household. A set of prices and the information on income and wealth enable us to distinguish combinations of goods and services that are available to a household from those that are not. The combinations of goods and services that are available to a household in a given time period are referred to as the *opportunity set,* and its size is defined by the household's *budget constraint.* To illustrate, consider a household with a known income, I, that purchases only two goods, X and Y, which sell for known prices P_x and P_y. We can write the budget constraint as follows

$$P_x \cdot X + P_y \cdot Y \leq I$$

where X is the amount of X purchased, and Y is the amount of Y purchased.

This statement says simply that the amount spent on X (which is $Px \cdot X$) added to the amount spend on Y (which is $Py \cdot Y$), must be *less than or equal to* income.

A budget constraint and opportunity set are illustrated geographically in Figure 4.1. In this diagram, each point represents a combination of X and Y. Point A, for example, represents X_a units of X and Y_a units of Y. The budget constant itself, line segment DE, shows all the combinations for X and Y that the household could buy if it spent all of its income. Consider point D. If the household chose D, it would be spending its entire income on Y and nothing on X; the total amount of Y that it could buy would be I/P_y. Similarly if it spent its entire income on X and nothing on Y, the total amount of X it could buy would be I/P_x.

All points below and to the left of the budget constraint are available and make up the opportunity set. It is represented by the shaded area in Figure 4.1. Point A is available, but the household does not spend its entire income. Point B is available, and the household spends its entire income. Point C is not available; that combination of X and Y (X_c and Y_c) cannot be purchased with an income of I.

The point of this exercise is to show that households face a *constrained choice problem.* The constraint is defined by income and prices. Within the limits imposed, households are free to choose, and the ultimate choice depends on the household's tastes or preferences. Obviously, the range of goods and services in society is large, and no two people have the same tastes or preferences.

The theory of household behavior is an attempt to derive some logi-

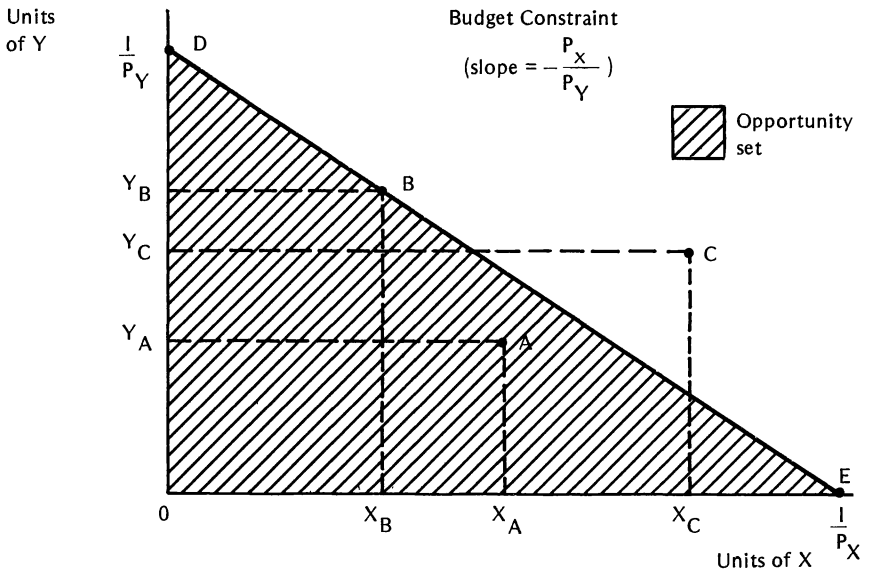


Figure 4.1. *Budget constraint and opportunity set for a household consuming two goods.*

cal propositions about the way households are likely to behave in response to changes in budget constraints. Tax policy can affect the opportunity sets facing households. Some taxes, such as the income tax, may affect spendable income directly. Others such as excise taxes may end up changing prices. Sometimes a tax may end up changing both prices and incomes. A corporate tax may reduce profits and thus the incomes of owners of firms; it may also cause the price of corporate products to rise or wages of workers to fall.

Income and Substitution Effects

Keeping in mind the constrained choice problem described above, consider the likely response of a household to a decline in the price of some product *X ceteris parabus*. How might a household currently consuming many goods be likely to respond to a fall in the price of one, when income, preferences, and all other prices remain unchanged?

Following the price decline the household faces a new budget constraint and may change its final choice of all goods and services. A decline in the price of gasoline may affect how much gasoline the household purchases but also the kind of car it buys, how much and in what ways it travels, and indirectly how many movies it sees this month.

Households are affected by price changes in two ways. First, if we

assume that households confine their choices to goods that improve their well-being ("more is better"), then a decline in the price of any good *ceteris parabus* makes the household unequivocally better off. If they buy the exact same amount of every good and service, they will have income left over. That extra income may be spent on X itself or on other goods. The change in consumption of X due to this improvement in well-being is called *the income effect* of the price change.

Second, when a product price falls, that product becomes cheaper in a relative sense; that is, relative to potential "substitutes" the good becomes more attractive. Even if the household's welfare were not improved by the fall in price, it may shift its purchasing patterns away from substitutes toward X. The change in the consumption of X because it is now relatively cheaper than potential substitutes is called *the substitution effect.*

Consider again a significant decline in the price of gasoline. This has actually occurred in *real terms* for most families in the United States since 1980: Incomes and other prices have risen faster than gasoline prices. With lower prices it is less costly to drive. I may visit my mother across the state more frequently and consequently go to the movies less frequently—partially because I am better off and can afford to travel more (*income effect*) but also because visiting my mother is now relatively cheaper than movies (*substitution effect*).

Household Choice and Economic Efficiency

Viewing the household choice problem in this way reveals much about the criterion of *economic efficiency*. Recall that a change is said to be efficient if those who benefit from it gain more than it would take to fully compensate those who lose. When we say that some people benefit, we really mean that they *themselves* determine that they are better off and by how much. The issue becomes how is it possible to measure such gains and losses?

If households are indeed free to choose within the constraints imposed by prices and incomes, their actual behavior reveals much about how they value various things being produced. If I decided to buy a cheeseburger for $2.50, I've "revealed" that it is worth *at least* $2.50 to me. The cost or value to me is really what I sacrifice by buying the item. By purchasing the cheeseburger I am giving up or forgoing the opportunity to buy other things with that $2.50. There is an *opportunity cost* implicit in my choice.

This point can be illustrated with a simple market demand curve like the one in Figure 4.2. Suppose that a group of sophisticated econometricians had used real data to estimate the demand curve in

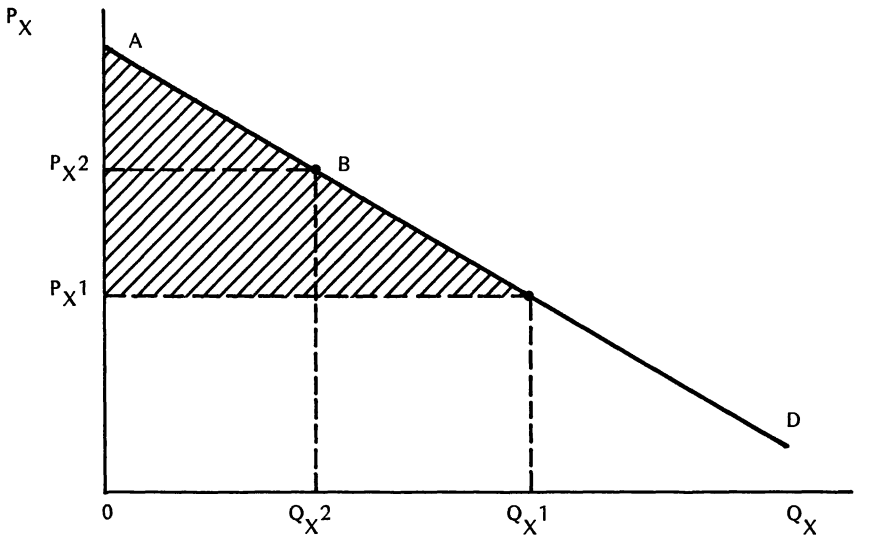


Figure 4.2. *Market demand, revealed preferences, and consumer surplus.*

the diagram. If the current market price is P_x^1, the demand curve shows that consumers will purchase Q_x^1. There is only one price in the market, and the demand curve tells us how much households would buy if they could purchase all they want at the posted price. Anyone who values a unit of X higher than P_x^1 will buy it; anyone who does not value it that highly will not. Recall people are free to choose whether to buy X and how much.

Notice that some people may value X at more than P_x^1. If, for example, price were P_x^2, Q_x^2 would be sold. If the market price turns out to be P_x^1, those people get a *surplus*. They value X at P_x^2 or higher, but only have to pay P_x^1. The total value of this *consumer surplus* is roughly equal to the area of the shaded triangle. To see this, think about offering X to consumers at successively lower prices; those near point A on the demand curve get a large surplus; those near point B get a smaller surplus. Since consumer behavior can be observed, the value of this surplus can be approximated.

The fact that such values are potentially measurable will be useful later on for making judgments about the efficiency of various tax policy alternatives. When a tax system is changed, prices adjust and consumers shift their buying patterns. Measuring the gains and losses associated with such changes will help in evaluation.

An important point to remember is that the market forces households to reveal their preferences because consumption is *contingent* on

payment. The theory of household behavior is sometimes called *revealed preference theory*: I don't get the cheeseburger unless I pay for it; thus, if I buy it, I reveal that it is worth to me at least the price I have paid. We will see later that there is a class of goods where this mechanism fails. If the benefits of a good are collective, it may be impossible to exclude people from enjoying them for nonpayment; the contingency is gone.

An example is reducing air pollution or providing police protection to a community. Residents of a particular community breathe its air, and if its crime rate falls, they have a reduced chance of becoming a victim. Both statements apply to residents whether or not they have paid for clean air or protection from crime. It will be argued later on that those kinds of goods, called *public goods,* will not be produced adequately by private firms.

Household Choice in Input Markets

Thus far, we have assumed that income is known or given. In fact, the income that a household receives depends, at least in part, on choices made by the household in input markets. Those who derive their income from working, for example, must decide whether to work, at what kind of a job, and whether to work full- or part-time.

The basic tradeoff involved in the decision to work and how much to work is the wages paid (and what it will buy) against the value of time spent if one does not work. The wage available to any given worker depends on the characteristics of the job and the skills of the worker. The decision process envisioned by economic theory is very similar to that described above.

Consider, for example, an individual facing a fixed wage of $\$W$, where work at that wage is his only source of income. If we define the alternative to work as *leisure* and endow everyone with twenty-four hours of time per day, Figure 4.3 illustrates the constraint within which preferences operate. If the individual decides to work eight hours (supply eight hours of labor), he consumes sixteen hours of leisure and earns an income of $\$(8 \times W)$.

When wage rates change, workers are affected in two ways. Consider, for example, an increase in the net wage rate that might result from an income tax cut. First, households are "better off," since by supplying the same amount of labor they earn a higher income. A higher income might also allow them to consume more of all goods: If we include leisure as a good, and it is a "normal" good, households may buy more of it. "Buying" more leisure is done by simply supplying *less* la-

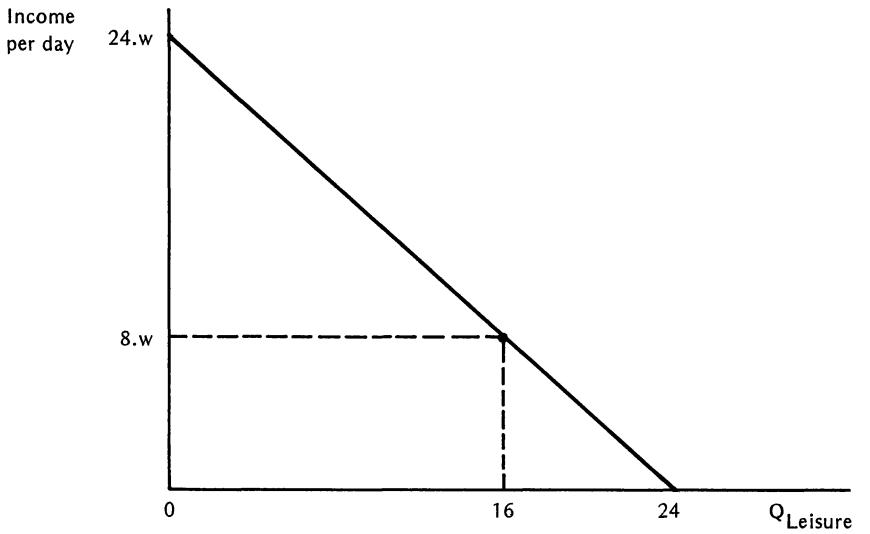


Figure 4.3. *The labor-leisure choice.*

bor, and its "price" is the forgone wage. This is called the *income effect* of the wage increase.

There is also potentially a *substitution effect.* When wages rise, leisure becomes more expensive relative to other goods, and individuals may substitute other goods for it. That implies buying *less* leisure and working more.

Note that in the labor market the income and substitution effects work in opposite directions if leisure is a normal good. The income effect of a wage increase implies buying more leisure and working *less*; the substitution effect implies buying less leisure and working *more.* Whether households will, in aggregate, supply more labor when taxes are cut will thus depend on the relative strength of income and substitution effects.

Review Questions and Exercises

1. Assume that as a result of the recent outbreak of hijackings and bombings, desire to fly diminishes significantly. Describe how you might expect the air travel market to react. What might happen to the price of airline tickets? Explain consumers' reactions to any price changes in terms of income and substitution effects.
2. Assume that consumers in Lumpland buy only two goods: X (a necessity) and Y (a luxury). Suppose Ms. A has an income of

$1,000 per year, the price of X is $5, and the price of Y is $10.

a. Sketch A's budget constraint.

b. Shade in A's opportunity set.

c. If A were relatively poor for Lumpland, where might you expect her to choose on the budget constraint?

d. Suppose that the price of Y were to fall from $10 to $5. Sketch the new budget constraint. What happens to the size and position of the opportunity set?

e. Suppose that prior to the price decrease, a tax of $2,010 had been placed on luxury goods. Assume that the tax is paid by the consumer. Sketch the new budget constraint facing Ms. A.

3. Sketch the income-leisure budget constraint facing an individual with:

a. A twenty-four-hour endowment of time daily;

b. Fifty dollars in property income per day (received regardless of work effort);

c. A job that requires a minimum of eight hours of work per day and that pays a wage of $\$W$ per hours, plus time and a half for all work over eight hours (1.5 x $\$W$);

d. No other work opportunities.

Note, All of this should be embodied in a single budget constraint.

5

Theory of the Firm: Profits and Production

Thus far we have discussed the behavior of households and the factors that influence demand in output markets and supply in input markets. We now turn to the other side of the system to explore the behavior of firms. Chapter 2 described business firms as economic agents that engage in production in order to earn profit: They supply outputs and demand factors of production in input markets.

We continue in this chapter to examine the operation of competitive markets. That is, we assume that there are many firms in each market, each small relative to the whole. It follows that each firm is a price-taker in both input and output markets; product price and factor prices are taken as given by individual firms. Again, this assumption will be relaxed below.

In the model to be developed, individual firms make three decisions simultaneously. Firms must choose

1. The quantity of output supply;
2. The quantities of each input to demand;
3. The technique to use in production.

Any two of these decisions determine the third. For example, once I have decided how much to produce and how to produce it, my input requirements are determined. Similarly, given a technique of production, any set of input quantities determines the amount of output that can be produced. In order to make these decisions, firms use three sets of data:

1. Input prices;
2. Techniques of production available;
3. The price of the firm's output.

The first two—input prices and available technologies—determine costs of production. Output price determines potential revenue.

This chapter begins with the nature of production and the notion of *technology of production*; the rest of the chapter explores the three decisions that must be made by producers and the factors that influence those decisions. A major goal of the chapter is to derive the supply curve of a single competitive firm in the short run.

Production and Production Technology

Production is a process of transformation in which inputs are combined, processed, and made into outputs. It is a process that takes many forms. General Motors uses land, buildings, labor of many different varieties, steel, rubber, and so forth and produces automobiles. The Boston Symphony Orchestra uses land, a beautiful building, highly skilled labor, musical instruments (capital), and so forth and produces performances.

Although we will focus in this chapter on private firms that produce for a profit, the process of production is not confined to them. Governments are also productive agents; they combine land, labor, and capital to produce desired public services such as national defense, police protection, and a system of laws. Households often act as productive agents as well. When I drive my car to work, I am combining my labor with capital to produce transportation services.

Production technology relates inputs to outputs. To produce any given service or good requires specific quantities of inputs. Just as a loaf of bread requires some amounts of water, flour, and yeast, as well as an oven and electricity, a trip from downtown New York to Newark, New Jersey, can be produced with a taxi cab, forty-five minutes of labor time, some volume of gasoline, and so forth.

Most outputs can be produced using a number of different combinations of inputs—that is, by using a number of alternative techniques. For example, tearing down an old building and clearing a lot to create a park can be done several ways. Five hundred men and women with small hammers could descend on it and carry the pieces away by hand; that would be a *labor-intensive* technology. On the other hand, the same park could be produced with two persons, a wrecking crane, a steam shovel, and a dump truck; that would be a *capital-intensive* technology.

Similarly, there are several combinations of inputs that may be used

to transport people from New York to New Jersey. A subway train can carry many people simultaneously; the amount of capital relative to labor in urban mass transit is very high. A taxi cab uses a much lower ratio of capital to labor but produces the service at a much higher cost per passenger trip.

Choosing the most appropriate technology is one of several important decisions facing firms. Presumably, firms choose the technology that minimizes the cost of production. In an economy with a plentiful supply of inexpensive labor and not a lot of capital, that will no doubt be a labor-intensive technique; in an economy with high wages there will be an incentive for firms to "substitute" away from labor and use more capital-intensive (labor-saving) techniques. We argue below that tax policy may influence the choice of technique with important consequences for the economy.

Production Functions

When the relationship between inputs and outputs (the technology of production) is expressed mathematically, it is called a *production function*. Table 5.1 and Figure 5.1 illustrate a simple production function for a small sandwich shop. Most of the sandwiches made are grilled, and the shop contains only one grill, which can easily accommodate only two people. When a third is added there is crowding, but with careful use of space more sandwiches can be produced. When the fourth and fifth workers are added they can move onto the grill while the others are wrapping and so forth, but then the first three must wait to get back on the grill. Hiring a sixth worker adds no output at all.

Table 5.1
Production Function for Sandwiches

Labor Units (employees)	*Total Product (sandwiches)*	*Marginal Product*
0	0	—
1	15	15
2	30	15
3	40	10
4	45	5
5	47	2
6	47	0

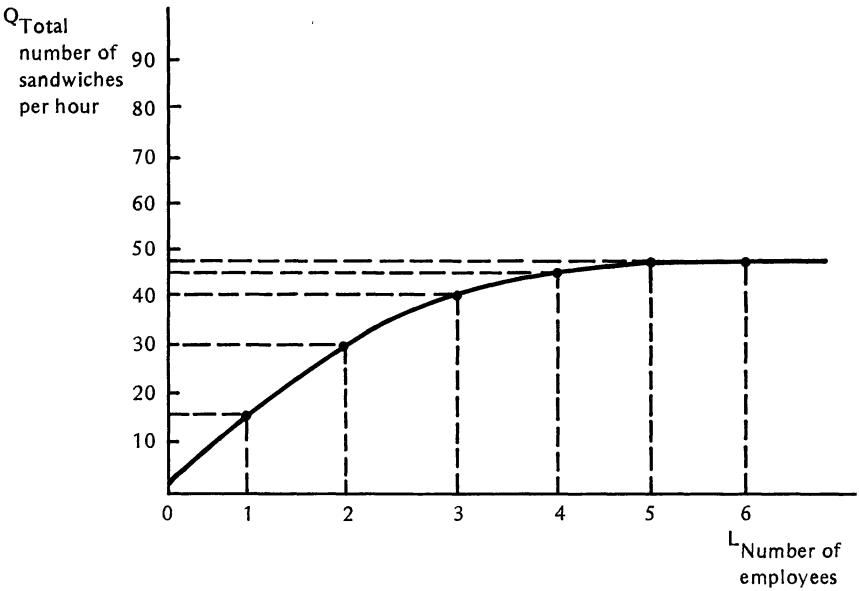


Figure 5.1. *Production function for sandwiches.*

Marginal product is defined as the additional output that can be produced by hiring an additional unit of input, holding all other inputs constant. Thus, from Table 5.1 the marginal product of the first unit of labor is fifteen sandwiches; the marginal product of the second is fifteen, the third, ten and so forth. Notice that the marginal product of the sixth is zero.

Figure 5.2 is a graph of the marginal product curve. Geometrically, it is a graph of the *slope* of the production function.

The *law of diminishing returns* states that when additional units of a variable input are added to fixed inputs (such as the building and grill), the marginal product of the variable input will decline. David Ricardo, a nineteenth-century British economist, was the first to point out diminishing returns in agriculture. With a given area of land, successive "doses" of labor and capital will yield smaller and smaller increases in crop output.

Diminishing returns can be seen in many productive activities. Consider, for example, an independent accountant who works primarily for private citizens preparing tax returns. As he adds more and more clients, he must work later and later into the evening. There is some point at which the additional hours spent yield little or no product. Here the "fixed" factor of production is the worker himself with a limited number of available hours during the day.

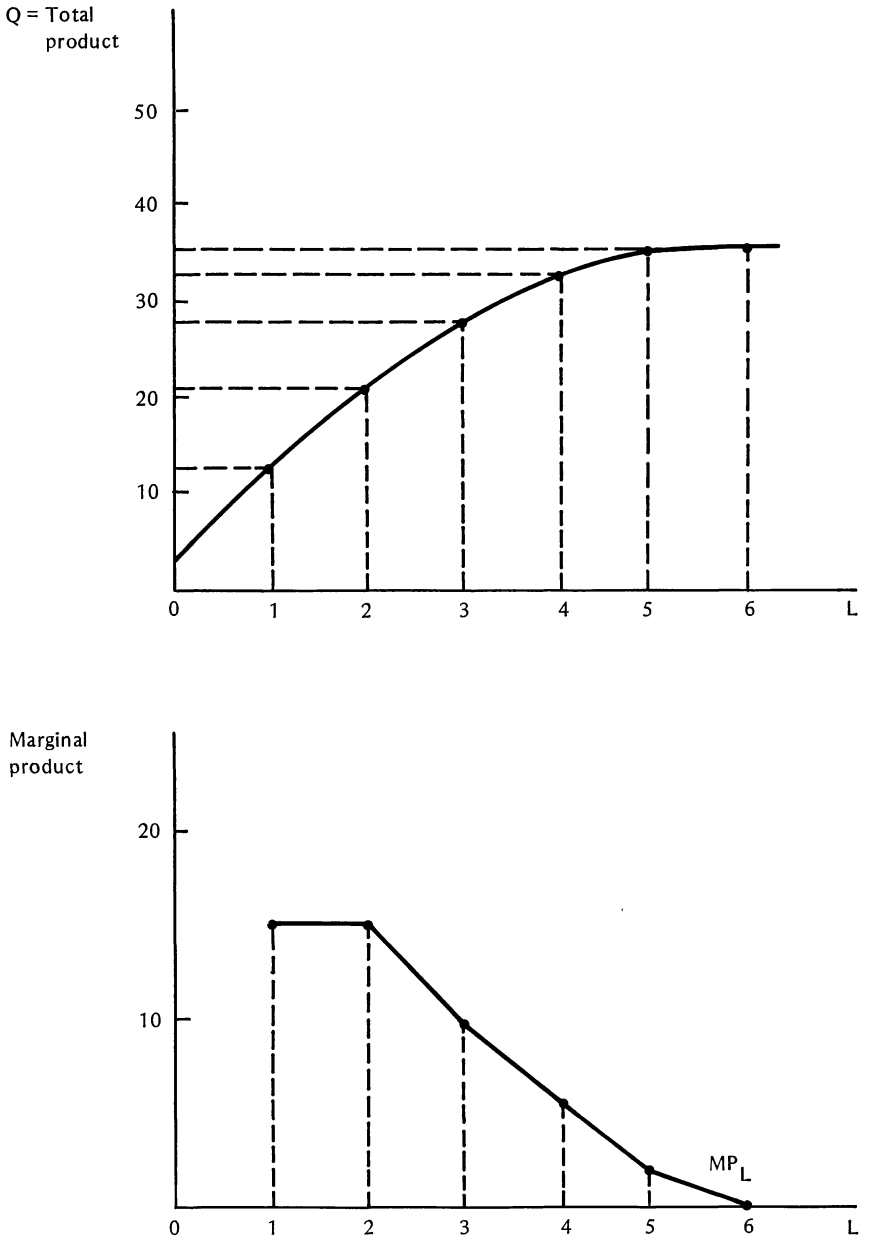


Figure 5.2. *Derivation of marginal product from a production function.*

Production functions may also represent more complex relationships between inputs and outputs. The production function below is one that expresses a relationship between output and two variable inputs:

$$Q = \sqrt{K} \cdot L = K^{1/2} \cdot L^{1/2}$$

Table 5.2 presents some numbers derived from this production function. They illustrate that several combinations of inputs can be used to produce each level of output.

This particular mathematical relationship also exhibits diminishing returns. Table 5.3 shows that if units of labor are added to twelve units of capital, each successive unit of labor produces a smaller amount of additional product.

Using actual data on firms of various sizes and at different points in time, economists have actually estimated production functions that correspond closely to measured inputs and outputs. As an example, the following function was estimated for paper production in India (see Edwin Mansfield, *Microeconomics: Theory and Application,* 5th ed., W.W. Norton, 1985, p. 175):

$$Q = A \cdot K^{.45} \cdot L^{.64}$$

Profits and Costs

It is assumed that the objective of firms is profit and that decisions are made in order to maximize profit. In simplest terms profits are the differences between revenues and costs:

Table 5.2
Combinations of Inputs That Can Be Used to Produce Six Units of Output ($Q = \sqrt{K \cdot L}$)

Q	*K*	*L*
6	36	1
6	18	2
6	12	3
6	6	6
6	3	12
6	2	18
6	1	36

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

Revenue are simply receipts from a sale where the number of units sold (q) times market price per unit (p). The economic definition of cost, however, differs slightly from "costs" as an accountant would calculate them. The most important difference is that economists add *a normal rate of profit* to costs. A normal rate of profit is the rate that is just sufficient to keep owners or investors interested in the firm.

From the standpoint of firm managers, this represents the *real cost of capital.* If the firm is operating with fairly steady profits and little risk, one would expect the normal profit rate to be close to the short-term interest rate—that is, if I can earn a return of 10 percent by buying a "safe" bond, I would put money into a "safe" firm only if the yield were slightly greater than 10 percent.

The implication of adding a normal profit rate to cost is that a firm earning a normal profit rate, earns no *economic profits.* Since normal profits are included in costs as the cost of capital, *economic profits* are profits *over and above* normal; they are sometimes called *excess profits.*

Choice of Technology and Costs of Production

The assumption that firms will maximize profits implies that they will minimize costs of production. Since firms have a number of production techniques to choose among, it follows that they will pick the technique that produces output at least cost. Clearly the choice will depend on the current market prices of inputs or factors of production.

Table 5.3
Diminishing Marginal Product of Labor $Q = \sqrt{K \cdot L}$

K	L	Q	*Marginal Product* ($\triangle Q/\triangle L$)
12	0	0	—
12	1	3.46	3.46
12	2	4.90	1.44
12	3	6.00	1.10
12	4	6.93	.93
12	5	7.75	.82
12	6	8.49	.74
12	7	9.17	.68

To illustrate consider the three techniques in Table 5.4. In the table, cost of production is calculated under two alternative sets of factor prices. Where labor and capital each cost $1 per unit, technique *B* at $14 per unit is the one chosen. However, if capital costs were to rise to $2 per unit technique *C* would be chosen; the firm would *substitute* capital for labor in response to the factor price change. It would use a more labor-intensive technique.

In Chapter 6, we will explore costs of production in more detail. The purpose of the discussion here is to highlight an important point: When a firm makes calculations comparing revenues and costs, determining cost of production requires analysis and involves considering factor prices as well as alternative techniques of production. Cost functions or cost curves are mathematical representations of the relationship between output and cost. Each point on a cost curve represents the lowest-possible cost at which that level of output can be produced.

Long Run Versus Short Run

When a tax is imposed or a tax system is changed, the immediate reaction of households and firms may differ from behavioral changes that take shape over time. The same can be said of household and firm adjustments to any change in the economic environment.

Consider, for example, the response of U.S. industry to the dramatic increases in crude oil prices in 1973 and 1974. Since most firms had, in a sense, built inexpensive oil into their operations, most had no choice but to continue to buy. Initially, demand for petroleum seemed to be very inelastic with respect to price.

Over a period of years, however, as new plants were built and old ones were modernized, firms adopted more energy-efficient techniques. Similarly, households responded over time by demanding more fuel-

Table 5.4
Choice of Technique to Minimize Cost

	Input Requirements Per Unit of Output		*Cost per Unit of Output*	
Technique	*Capital*	*Labor*	$P_L = \$1$ $P_K = \$1$	$P_L = \$1$ $P_K = \$2$
A	10	5	$15	$25
B	7	7	[$14]	$21
C	4	12	$16	[$20]

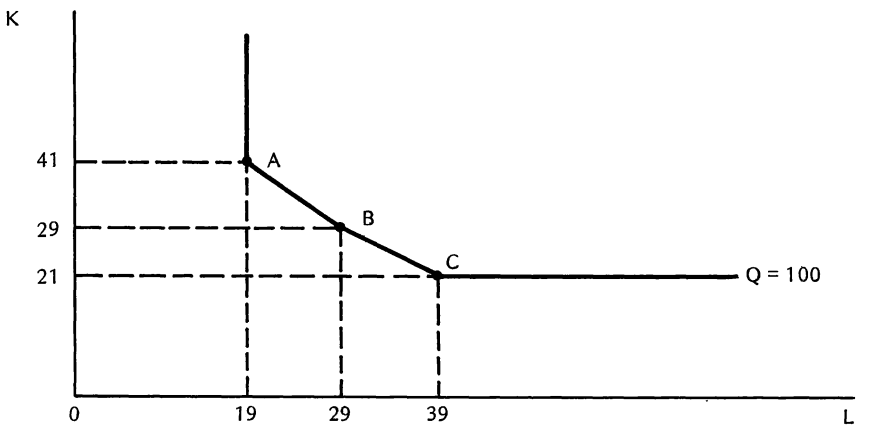
efficient automobiles and by insulating their homes. Ultimately, the demand for petroleum was significantly reduced; demand in the long run was more elastic.

Because the character of immediate response differs from long-run adjustment, it is useful to loosely define two time periods: *the long run* and *the short run.* Two assumptions define the short run: fixed scale and no entry or exit. First, existing firms have some fixed factor of production; they are locked into their current scale of operations. Second, new firms cannot enter an industry, and existing firms cannot exit. Firms may curtail operations, but they are still locked into some costs even if they will ultimately go out of business.

In the long run, no factors of production are fixed. Firms may double or triple capacity or go completely out of business. Presumably, supernormal profits will attract new firms.

Review Questions and Exercises

1. The diagram (an isoquant) below shows the various combinations of capital and labor that can be used to produce 100 units of output:



Consider the techniques represented by points *A, B,* and *C.* If the price of labor were $2 and the price of capital were $2, and the firm decided to produce 100 units of output, which technique would be employed? If a payroll tax were imposed that pushed the price of labor to $3 but left the price of capital at $2, which technique would be chosen? If a profits tax pushed the price of

capital up to $3 but left the price of labor at $2, which would be chosen?

2. Assume that widgets can be produced using two different techniques, *A* and *B*. The following table provides data on the total input requirement of each at four different output levels:

	Q = 1		*Q* = 2		*Q* = 3		*Q* = 4	
	K	*L*	*K*	*L*	*K*	*L*	*K*	*L*
A	5	2	8	4	11	5	15	5
B	3	3	6	5	8	7	11	10

   a. If labor is $2 per unit and capital is $3 per unit, what is the minimum cost of producing one widget? two? three? four?
   b. Graph cost of production as a function of output.
   c. How much did it cost to go from an output of one to an output of two? To go from an output of two to three? From three to four?

3. Do you think demand (supply) is likely to be more elastic in the short run or in the long run? Why? What factors are likely to influence long-run supply elasticity? Demand elasticity?
4. Suppose that it cost $500,000 to start an ice cream store ($500,000 includes all the equipment necessary to operate). Assume that annual revenues from operating the store would be $80,000 and that all costs (as an accountant would measure them) would be $50,000. The costs include rental of the store, labor, supplies, and full maintenance of the equipment. Full maintenance implies that if our owner went out of business, she could sell the equipment for the full $500,000. Suppose that the ice cream business is risky and that a rational capitalist would demand a 20 percent return on investment to enter. What is profit as an accountant would measure it if our capitalist bought the equipment with her own money? If you consider 20 percent a normal return to investment, how much economic profit would the enterprise earn? Explain your answer.

6

Theory of the Firm: Cost and Output Supply in Competitive Markets

The previous chapter focused on one of the simultaneous decisions that competitive firms must make: choice of technology. It should be recalled that technological choice is assumed to be made on the basis of cost: Firms pick the technique that minimizes the cost of producing the desired level of output, and that choice depends on factor prices. Cost of production, then, depends on *technologies available* and *factor prices*. This chapter argues that final output supply depends on *costs* and *output prices*.

Fixed and Variable Costs: The Short Run

You will recall that two conditions define the short run: (1) Existing firms have some "fixed" factor of production; and (2) new firms cannot enter an industry, and existing firms cannot exit, or go completely out of business. The notion of the short run has implications for costs.

First, some costs must be borne by each firm regardless of output; these are called *fixed costs*. They must be paid even if the firm shuts down its operations—that is, even if output is zero. If output is large, they do not change.

Sometimes fixed costs are called overhead. If you operate any kind of plant facilities, you must heat it or keep the pipes from freezing in the winter even if no production is taking place then. There may be contract obligations and an administrative staff to pay.

For some firms fixed costs represent a large portion of total costs. An electric company generally has a large generating plant, thousands of miles of distribution wires and poles, transformers, and so forth. Usually such a plant is financed by borrowing huge sums of money or issuing bonds to the public. The interest that must be paid on these funds represents a substantial part of the cost of the utility; it is a cost that does not depend on how much electricity is being produced.

Since fixed costs do not vary with output, a graph of fixed costs against output would be a simple horizontal line. In Table 6.1 and Figure 6.1 it is assumed that our firm has fixed costs of $1,000.

Average cost is total cost simply divided by the number of units of output. Average fixed cost is total fixed cost divided by output. For example, if our simple firm produced five units of output, average fixed cost would be $200 or $1,000 divided by five. A graph of average fixed cost would be a downward sloping curve (see Figure 6.1). It would assymptote to or approach zero as the quantity of output grew. For example, if output were 100,000 units, average fixed cost would equal only 1 cent per unit in our example.

A second implication of the short-run definition is that cost per unit will eventually rise with output. The assumption of a fixed factor of production implies that the firm is stuck at its current *scale of operation.* As a firm tries to increase its output, it will eventually be trapped by its current scale. It can hire more labor and materials (variable inputs), but *diminishing returns* operates.

The example of the sandwich shop discussed in Chapter 5 illustrates the principle. With a fixed grill capacity, more sandwiches can be produced with more laborers, but the marginal product of each successive labor unit applied declines as the grill gets crowded.

Costs that depend on the level of output chosen are called *variable costs.* Fixed costs and variable costs together make up total costs:

$$TC = TFC + TVC$$

Table 6.1
Short-Run Fixed Cost (Total and Average) of a Hypothetical Firm

Q	*TFC*	*AFC*
0	1,000	—
1	1,000	1,000
2	1,000	500
3	1,000	333
4	1,000	250
5	1,000	200

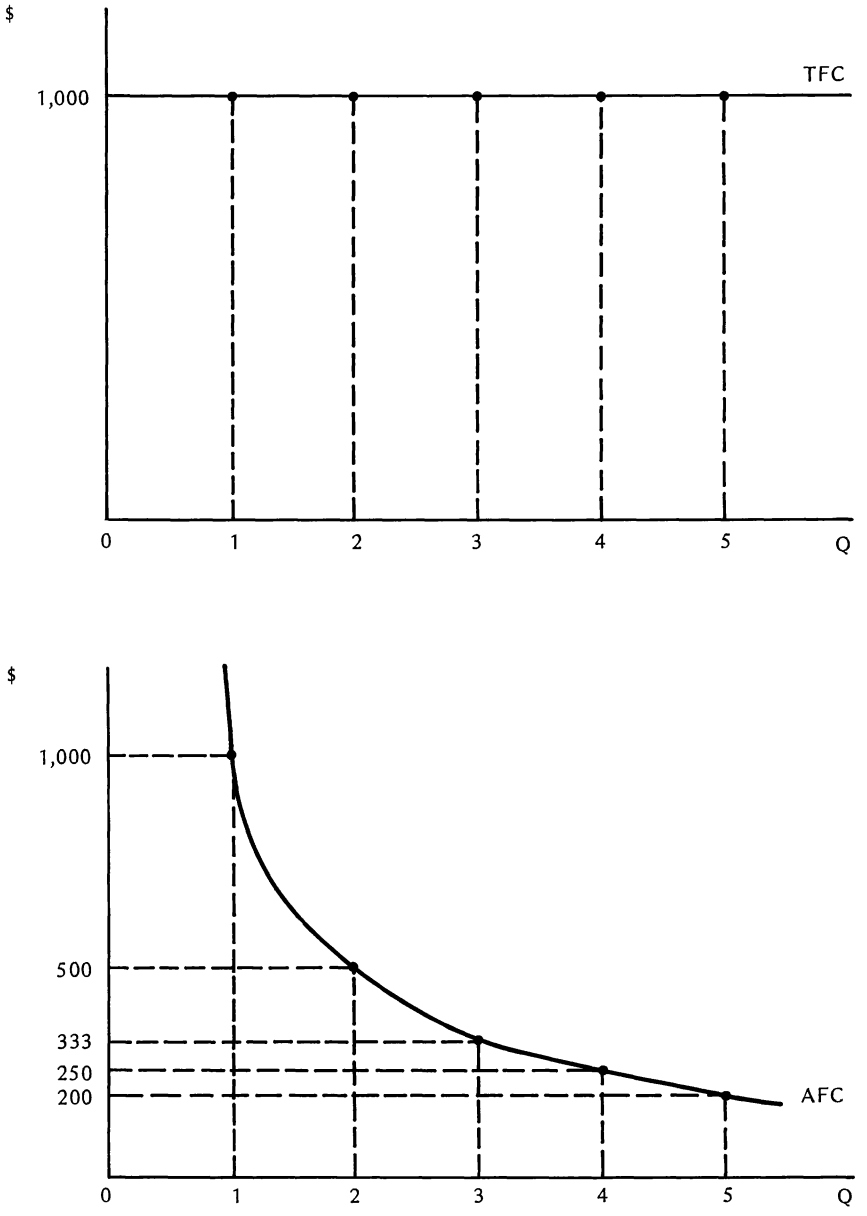


Figure 6.1. *Short-run fixed cost (total and average) of a hypothetical firm.*

Marginal Cost: The Short Run

The most important of all cost concepts is *marginal cost (MC).* Marginal cost is the increase in total cost caused by the production of one more unit. If the firm is producing 1,000 units of output and decides to raise output to 1,001, costs will rise. The increase, which amounts to the cost of the *one thousand first unit* is called marginal cost.

Notice that marginal costs are one way of looking at variable costs: Marginal costs *are* variable costs. In Table 6.2 producing the first unit of output increases total cost from $1,000 to $1,010 and total variable cost from $0 to $10. Marginal cost is thus $10.

The second unit raises total variable cost from $10 to $18. Marginal cost is thus $8 and so forth. Notice that marginal cost is only the cost of the unit in question itself. *Average cost* is the average cost of *all* the units being produced. Consider again Table 6.2. The marginal cost of the third unit is $6, since total variable cost rose from $18 to $24. The *average* variable cost of the first *three* is $8. The first costs $10, the second costs $8, and the third costs $6; the average is $8 or $24 divided by 3.

Notice that average cost and marginal cost are related in a very specific way. When marginal cost is *below* average, average cost will decline toward it. Think, for example, of test scores. If you have answered an *average* of 85 percent questions correctly on three exams and you then answer 75 percent correctly, your average will fall. In Table 6.2 the average cost of producing *two* units is $9. The marginal cost of the third is $6, and the average falls to $8. Similarly, when marginal cost is above average, average cost will increase toward it. If you

Table 6.2
Short-Run Costs of a Hypothetical Firm

Q	*TVC*	*MC*	*AVC*	*TFC*	*TC*
0	0	—	—	1,000	1,000
1	10	10	10	1,000	1,010
2	18	8	9	1,000	1,018
3	24	6	8	1,000	1,024
4	32	8	8	1,000	1,032
5	42	10	8.4	1,000	1,042
—	—	—	—	—	—
—	—	—	—	—	—
—	—	—	—	—	—
500	8,000	20	16	1,000	9,000

answered 95 percent correctly on your last test instead of 75 percent, your average would rise. In Table 6.2, the average cost of *four* units is \$8. The fifth unit costs \$10, and the average rises to \$8.40.

In the last section it was pointed out that the concept of the short run had implications for costs. The most important is that fixed scale of plant causes diminishing returns to variable inputs resulting in rising unit costs. This means that in the short run, marginal cost will eventually rise as output increases.

Graphing Total, Average, and Marginal Costs

Figure 6.2 shows a typical short-run variable cost curve that ultimately exhibits increasing marginal cost. The relationship between the total variable cost curve and the marginal cost curve is important. Since marginal cost is by definition the change in the variable (or total) cost resulting from a single unit increase in output, marginal cost is simply the numerical value of the slope of the total variable cost curves: $\Delta C/\Delta Q$. If the change in Q as we move along the quantity axis is *one* each time, then $\Delta C/\Delta Q = MC$. On the graph, the marginal cost of the second unit is \$8. The slope of the TVC curve between 1 and 2 units of output is $\Delta TVC = 8$ divided by $\Delta Q = 1$ which is 8, which is marginal cost.

Notice that the slope of the total variable cost curve falls (it gets flatter) up to point A. Correspondingly, the marginal cost curve declines. At twenty-nine units of output, the slope of the TVC curve gets steeper; marginal cost at that point begins to rise.

Average cost follows marginal cost, but it lags behind because it is the average of all previous units. Figure 6.3 shows average variable costs for our hypothetical firm following marginal cost down until forty-eight units of output are produced. Marginal cost begins to rise at twenty-nine units, but average cost does not begin to rise until marginal cost crosses it and rises above it. That occurs at forty-eight units. As a result, it is always true that marginal cost intersects average variable cost at its lowest or *minimum point.* Be sure you understand why.

Figure 6.4 adds fixed costs to the picture. Total cost is the simple addition of total fixed and total variable costs. This is shown in the top diagram where the same vertical distance (equal to FC) is simply added to TVC. If $TC = TVC + TFC$, it follows that $TC/Q = TVC/Q + TFC/Q$ or $ATC = AVC + AFC$. The difference between AVC and ATC is thus AFC, and that difference shrinks with output. Recall from Table 6.1 and Figure 6.1 that average fixed cost (AFC) falls steadily with output.

For exactly the same reasons that it intersects the minimum point of the AVC curve, marginal cost intersects ATC at its minimum point.

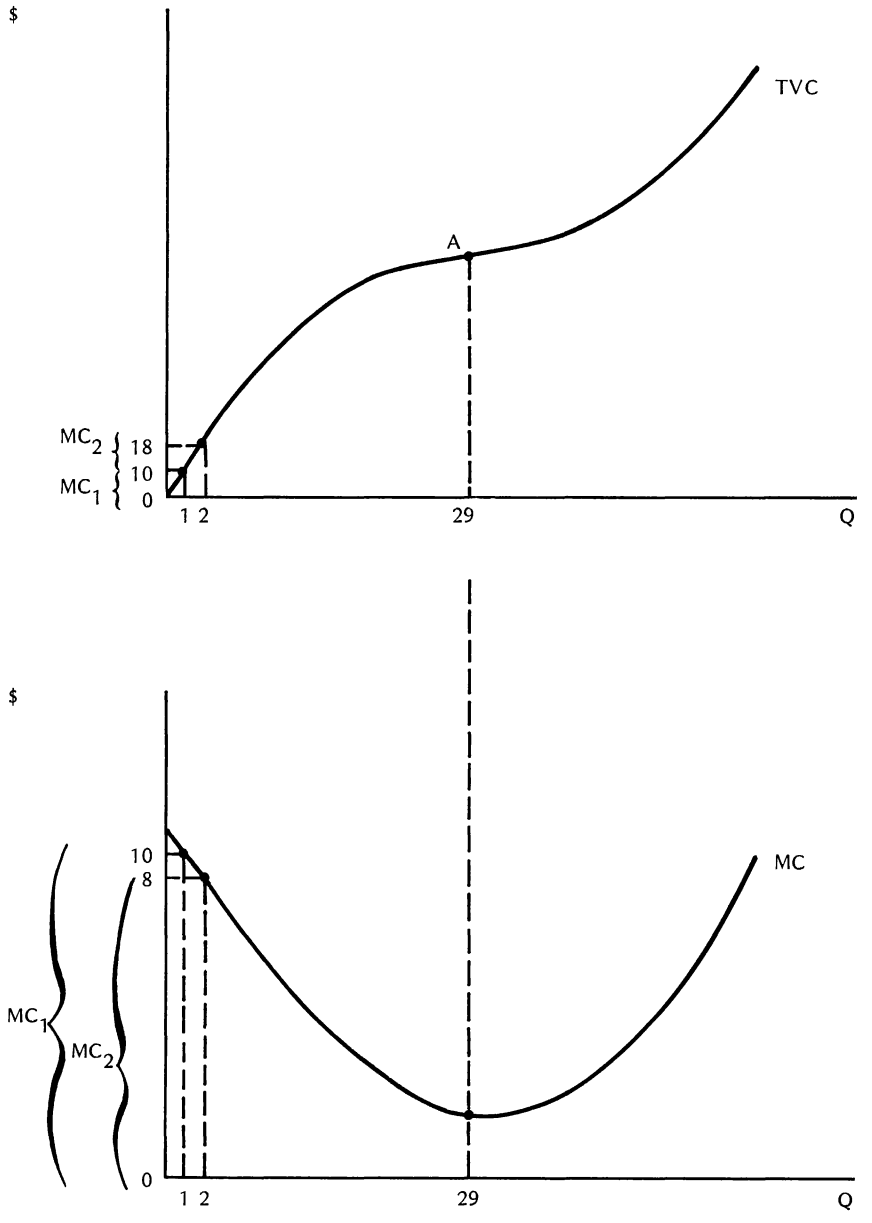


Figure 6.2. *Total variable cost and marginal cost for a typical firm.*

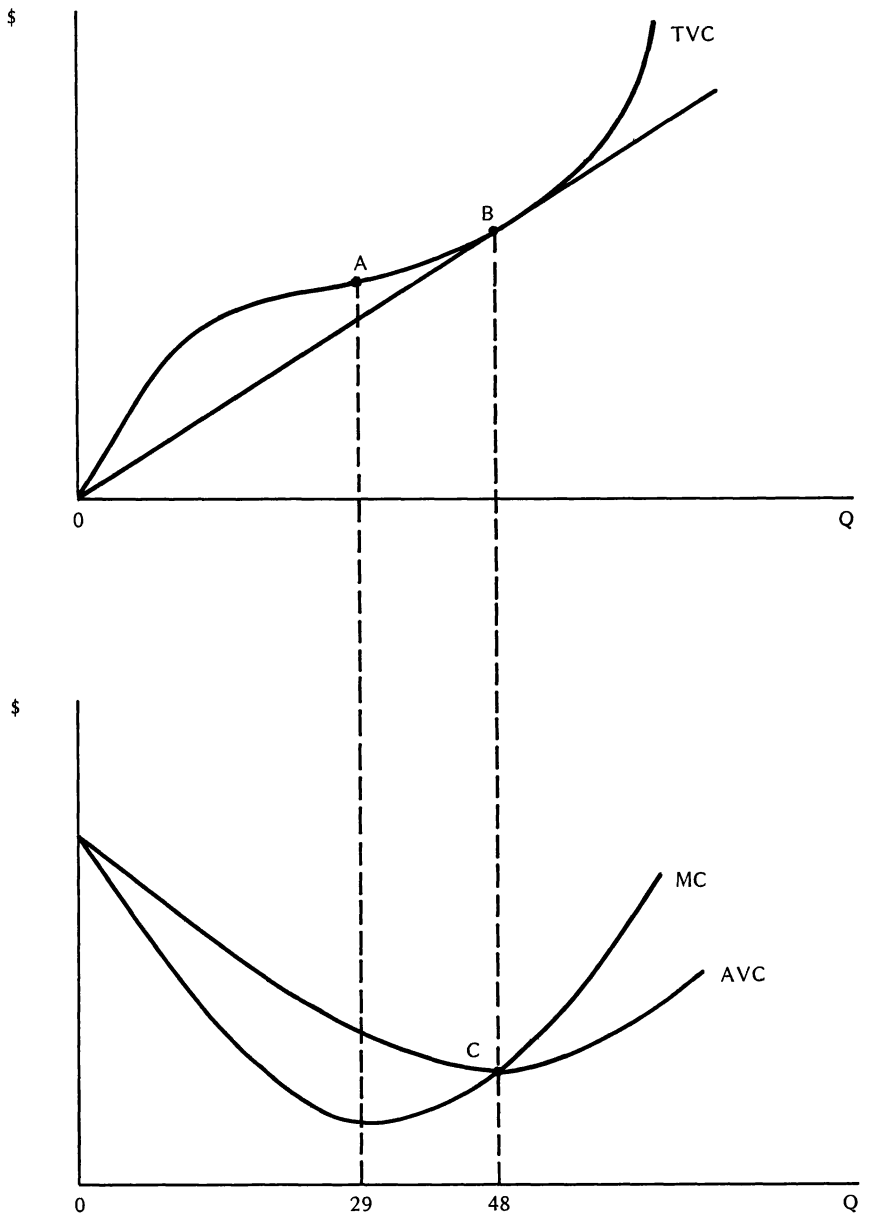


Figure 6.3. *More short-run costs.*

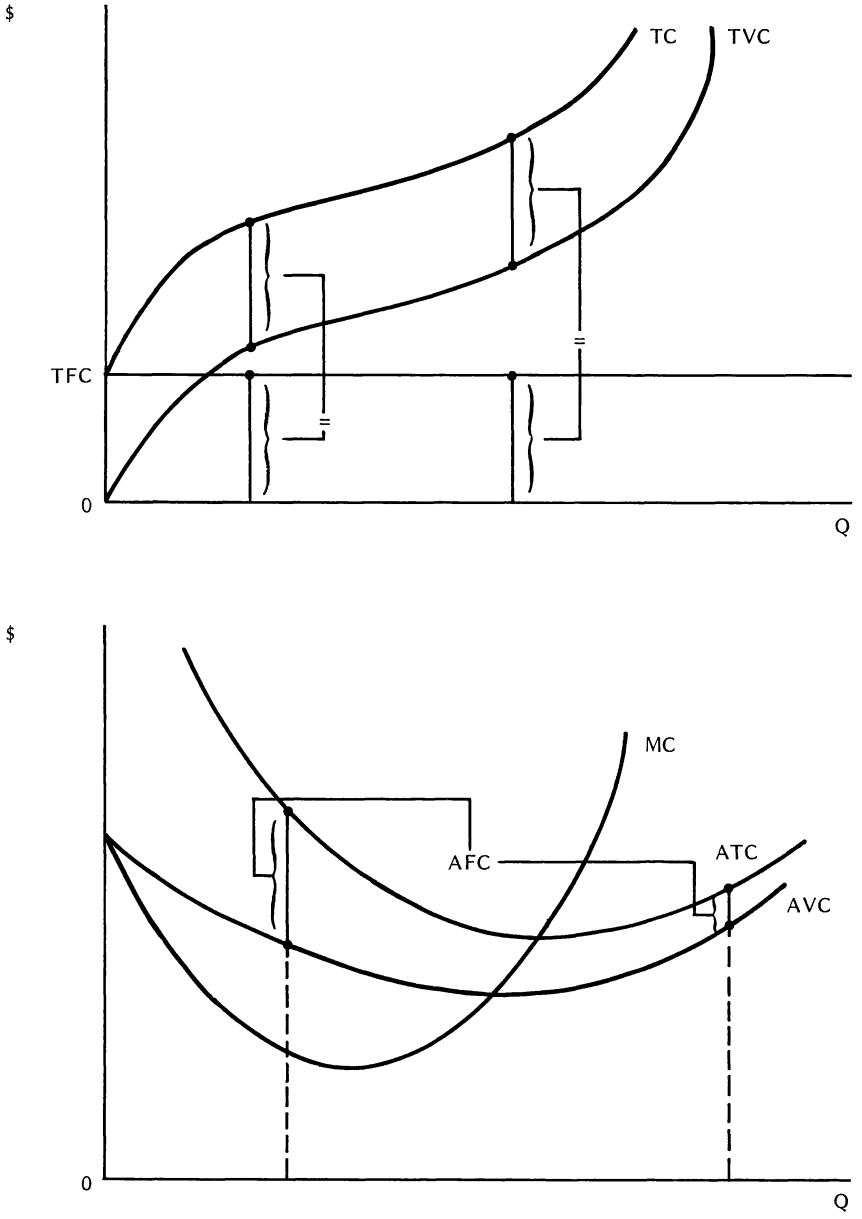


Figure 6.4. *Even more short-run costs.*

Since fixed costs do not change with output, marginal costs alone drag average total cost around; if *MC* is below *ATC*, *ATC* falls, and if *MC* is above *ATC*, *ATC* rises.

Marginal Cost as the "Supply Curve"

To derive the cost functions, firms take information on input prices and available techniques and choose the process that minimizes cost of production at each potential output level. Our typical firm must turn to its output market in order to make the final supply decision. Recall that we are assuming competitive behavior; each firm is small relative to the market and thus can have no influence on market price. The firm assumes that it can sell all it wants at the prevailing market price.

Recall also that profit is total revenue minus total cost and that total revenue from the sale product is simply $P \times Q$. *Marginal revenue* is defined as the additional revenue that a firm earns when it raises output by a single unit. In competition, where each unit of output sells for the same price, marginal revenue is simply equal to the prevailing market price.

Figure 6.5 presents the situation facing a typical competitive firm. The diagram shows a market equilibrium determined by the interaction of all suppliers and demanders. Each firm observes only the outcome: the market price.

The firm now uses all the information available to decide on the output level that would generate maximum profit. That is, it must choose the profit-maximizing quantity of output to supply. Careful thought will show that each firm will produce as long as the marginal revenue from an additional output exceeds marginal cost. If, indeed, raising output (or, in the planning stage, raising the target output) by one unit will generate revenue in excess of cost, the excess represents *added* profit. In the diagram of the typical firm in Figure 6.5, consider the firm's choice of output at price P_o. Recall that in competition *marginal revenue* is the prevailing market price. If I produce an added unit, I simply sell it for P_o, and revenues go up by P_o. In the diagram, at very low levels of output P_o is clearly greater than marginal cost, and the firm can add profit by raising output. Consider, for example, *MR* and *MC* at q_2. Raising output by a unit at q_2 will generate revenues of MR_2 (or P_o), but costs will rise only MC_2. Thus, an added unit of output *increases* profit by $(MR_2 - MC_2)$.

This logic continues until q_o. At q_o marginal cost rises above marginal revenue; added output will generate cost in excess of revenue. Thus, q_o, where $P_o = MC$, is the level of output that maximizes profit.

If market demand were to shift from D to D' in the market (see left

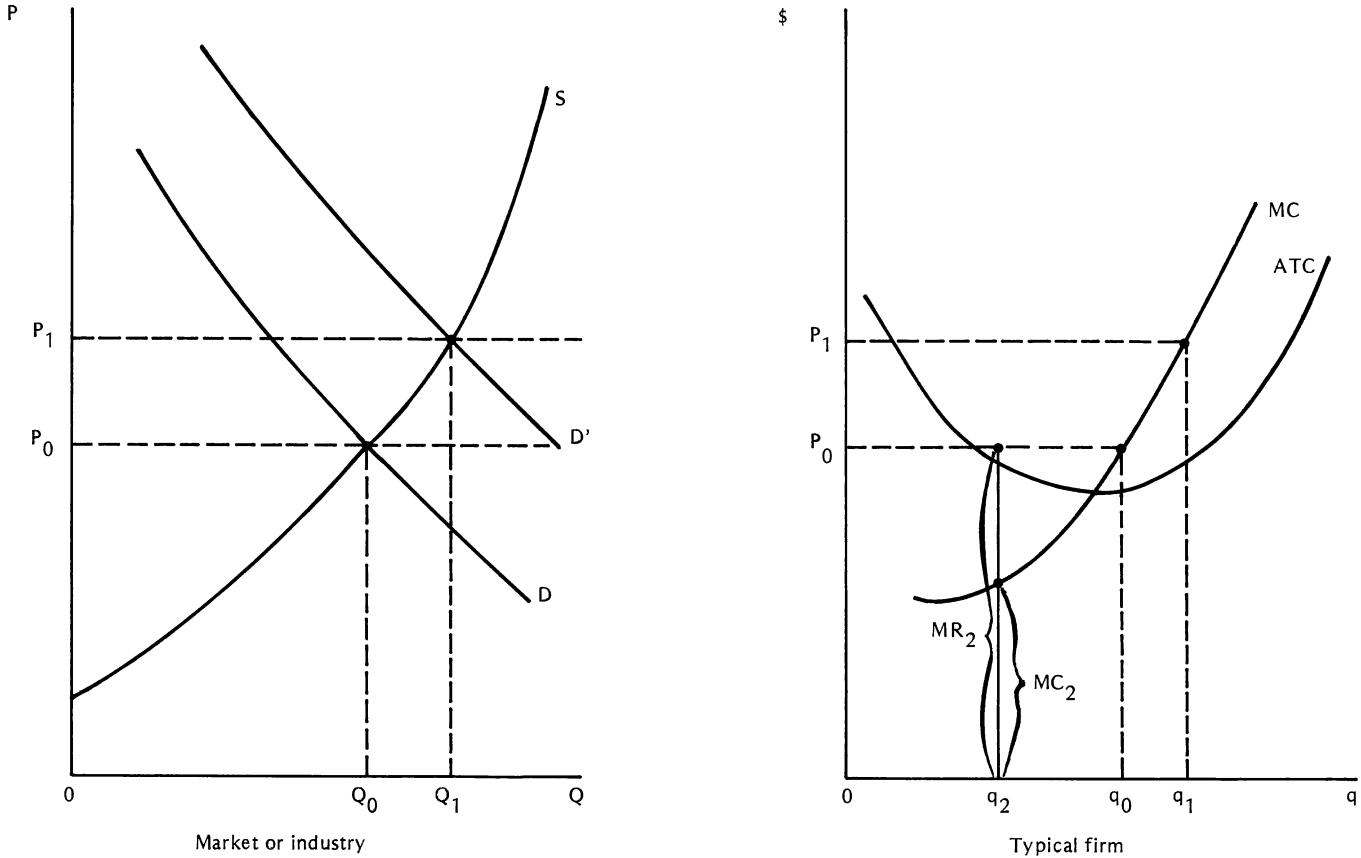


Figure 6.5. *Profit-maximizing output for a competitive firm.*

diagram in Figure 6.5) and price were to rise to P_1, the typical firm would find it advantageous to increase output. With the higher price (and thus marginal revenue) the firm can produce added profit by raising output to q_1. Again price is equal to marginal cost: At the profit maximizing level of output, $P_1 = MC$.

Notice that we have identified a curve—the marginal cost curve—that gives the profit maximizing output level at each market price. *Thus, the marginal cost curve is the supply curve of a competitive firm.* We will see in a moment, however, that there is one exception to this basic rule.

Figure 6.6 shows a firm earning economic profits in the short run. With a market price of P^* the firm will maximize profits by producing at q^*. At q^* total revenue (TR) is equal to P^* times q^*, which is equal to the area of rectangle P^*Aq^*O. Also, at q^* total cost (TC) is equal to average total cost (ATC) times q^*. (Since average total cost is total cost divided by q^* by definition, total cost is average cost *times* q^*.) Thus, total cost is the area $OCBq^*$. Profit is simply $TR - TC$, or the area of rectangle $ABCP^*$, which is shaded.

Recall that economic profits are profits *over and above* a normal return to owners of the firm. Normal profit is included in cost. In the situation above, there is an incentive for capital to "flow into" this industry in the longer run. Existing firms have an incentive to expand, and new firms are likely to enter the business to take advantage of the high profit rate.

Figure 6.7 shows a typical firm in an industry facing a lower market price, one that is insufficient to cover full costs. Again, the firm will continue to produce as long as marginal revenue (P^*) exceeds marginal cost of production (MC). Optimal output is thus q^*.

At q^*, however, there seems to be a problem. Total revenue (TR) is equal to P^* times q^* or the area of rectangle P^*Aq^*O. Total cost is equal to ATC times q^* or $OCBq^*$. Total cost is now greater than total revenue, indicating that the firm is showing a loss. Note that the term *loss* may or may not indicate losses in the accounting sense. A firm suffering economic losses may simply be earning profits *below* the normal profit rate.

Clearly, firms suffering losses have an incentive to get out of the business in the long run, but recall they are struck in the short run. They cannot exit, and they must pay fixed costs regardless of how much they produce. They can however "shut down," and this option must be explored: Should they continue to operate or not as they move toward long-run exit?

To answer this question the firm would simply compare the revenues and costs associated with producing and not producing. If the firm shuts down, it earns no revenue and incurs no *variable* costs but must pay *fixed* costs; losses are thus equal to fixed costs. Where it operates, it

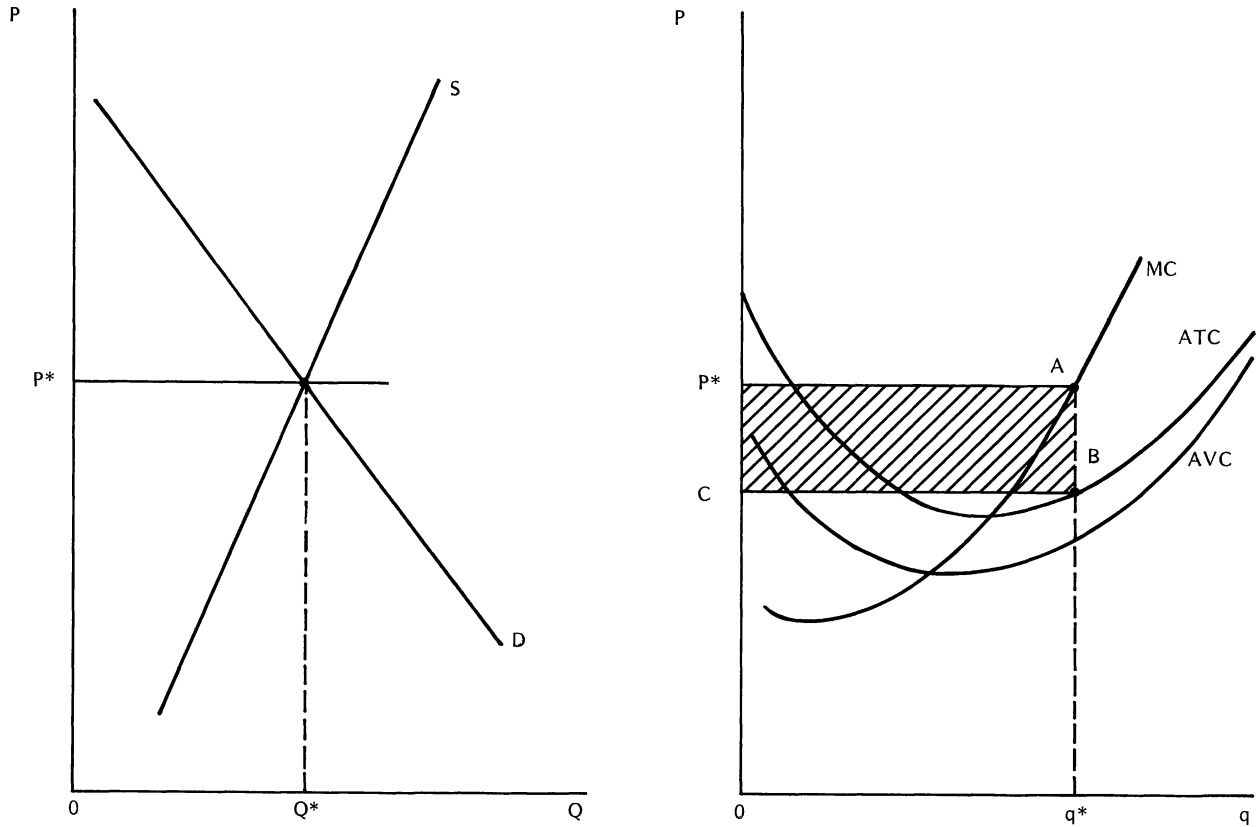


Figure 6.6. *Firm earning economic (excise) profits in the short run.*

earns revenues and incurs *variable* costs, and if those revenues exceed the variable cost incurred, the resulting *profit on operation* can be used to cover some of the fixed costs. (Recall that fixed costs must be paid whether or not the firm operates.)

If the firm in Figure 6.7 operates, it will do best by producing q^* units. At that output total revenues are P^* times q^* or the area of rectangle P^*Aq^*O. Total variable cost is equal to average variable cost times q^* or the area DEq^*O. Thus, revenues exceed variable costs by P^*AED, which is *profit on operation*. Thus, the firm *minimizes its losses* by operating in the short run and supplying q^* units of output.

Note that at q^* the difference between ATC and AVC is the length EB, which is average fixed cost (AFC). Average fixed cost times q^* is total fixed cost, which in the diagram is area $CBED$. This area is equal to the loss that would be suffered if the firm shut down. Profits on operation (area P^*AED) reduce them, and actual losses are only $CBAP^*$.

The result is that firms will continue to produce in the short run as long as price is sufficient to cover variable costs. If market price falls *below* variable costs, however, operating would only add to losses, and the firm will shut down. The bottom of the average variable cost curve is thus called the *shutdown point*. The supply curve of a competitive firm then is its marginal cost curve *above the average variable cost curve.* At any lower price output is zero.

It should be recalled that the market supply curve is the horizontal addition of the supply curves of all the firms in the industry. Thus, the market supply curves labeled S in Figures 6.5 through 6.7 are simply the horizontal sums of the marginal cost curves of all the firms in the industry.

Review Questions and Exercises

1. A firm's cost curves are given by the following table:

Q	TC	FC	VC	AVC	ATC	MC
0	100	100				
1	130	100				
2	150	100				
3	160	100				
4	172	100				
5	185	100				
6	210	100				
7	240	100				
8	280	100				
9	330	100				
10	390	100				

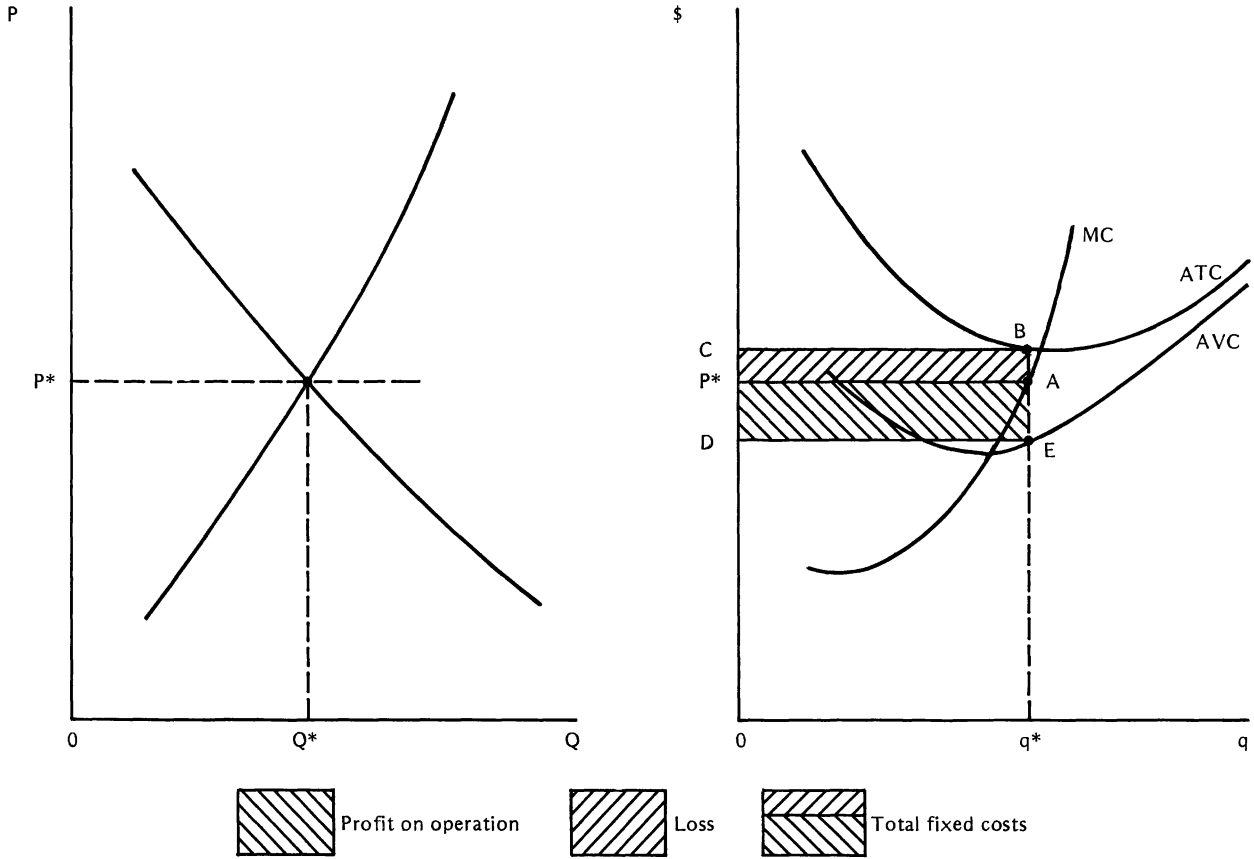
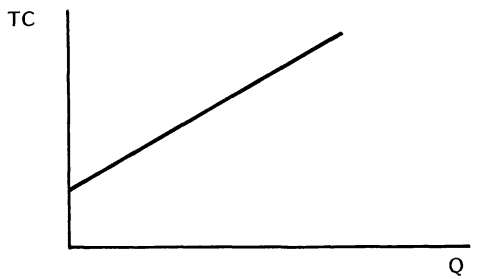


Figure 6.7. *Firm suffering economic losses but showing a profit on operation in the short run.*

a. Copy and complete the table.
b. Graph *ATC, AVC,* and *MC* on the same graph. What is the relationship between the *MC* curve and *ATC*? Between *MC* and *AVC*?
c. Suppose the market price is 30; how much will the firm produce in the short run? What are the profits? Show profits on the graph.
d. Suppose the market price is 50; how much will the firm produce in the short run? How much are the profits? Show them on the graph.
e. Suppose market price is 10. How much would the firm produce in the short run? What are the profits? Show profits on the graph.

2. The total cost curve for producing a book is given below. Draw the total fixed cost curve, the average fixed-cost curve, the *AVC* curve, the *ATC* curve, and the *MC* curve.



3. A 1973 Harvard graduate inherited her mother's printing company. The capital stock of the firm consists of three machines of various vintages all in excellent condition:

	Cost of Printing & Binding Per Book	*Maximum Total Capacity per Month*
Machine 1	$1.00	100 books
Machine 2	$2.00	200 books
Machine 3	$3.00	500 books

Assume that costs of printing and binding per book include *all* labor and materials including her own wages. Assume further that Mom signed a long-term contract (fifty years) with a service company to keep the machines in good repair for a fixed fee of $100 per month.

a. Derive the marginal cost curve of the firm.

b. Derive the total cost curve of the firm.

c. If she could sell all the books she wanted at $3 each, how many per month would she produce? What would her total revenues be? Total costs? Total profits?

7

Theory of the Firm: Demand in Competitive Input Markets

This chapter continues the discussion of decisions that competitive firms face in the short run. Recall that firms face three fundamental decisions: (1) which technology to employ in production, (2) how much to supply in output markets, and (3) how much to demand in each input market. In making these decisions, firms analyze three important sets of data: (1) input prices, (2) available technologies, (3) output price.

In the previous chapter, the focus was on outputs. The profit-maximizing level of output to be supplied depended on cost of production and its relation to the market price of product. The final decision involved comparison of revenues and costs from marginal units of output.

This chapter focuses on input demand, but it is important to understand that all three decisions are made simultaneously; in a sense they are simply different perspectives on the same decision process. Any two imply an answer to the third. For example, once a firm has chosen the least-cost technology and picked the profit-maximizing level of output, input requirements are determined.

Marginal Revenue Product

The *marginal revenue product* (*MRP*) of an input is the additional revenue a firm will earn by employing one additional unit of the input *ceteris parabis*. If a firm were to employ an additional unit of labor, *MRP* is

the revenue gained from the sale of the output produced by that unit of labor. Recall from Chapter 5 that the output produced by one additional unit of labor is called the *marginal product of labor.*

For a competitive firm then, marginal revenue product is simply the factor's marginal product times the price of output or the value of the factor's marginal product. For labor:

$$MRP_l = MP_l \times P_l$$

Table 7.1 reproduces the data from the hypothetical sandwich shop discussed in Chapter 5. Marginal product declined, since in the short run there was a fixed factor of production (grill capacity). As a result, marginal revenue product declines from $45 for the first worker hired to zero for the sixth.

Input demand depends on the relationship between the marginal revenue product of the input and its unit cost or price. For labor, unit price is the wage that is determined in the labor market. In the sandwich shop example, *daily wage* can be thought of as the marginal cost of a unit of labor. A profit-maximizing firm will hire an input as long as its marginal revenue product exceeds its price.

Suppose, for example, that sandwich makers were paid $28 per day for their services. The sandwich shop would hire three. The marginal revenue product of the third was $30 (more than $28), but the fourth would add only $15 in revenue.

Figure 7.1 is a diagram of a hypothetical firm facing a market wage rate, W^*. The diagram is perfectly analogous to Figure 6.6, in which the firm compares the marginal revenue (or price of a unit of *output*) with the marginal cost (*MC*) of a unit of *output*. Here the firm compares the marginal revenue of a unit of *input* (marginal revenue product) with the marginal cost of a unit of *input* (wage rate or input price).

Firms will hire up to the point that the wage is equal to labor's marginal revenue product:

$$w = MRP_l$$

The marginal revenue product curve, then, shows the profit-maximizing level of input demand at every potential input price. For a firm that employs only one variable factor of production, the MRP curve for that factor is the firm's demand curve.

Two Variable Factors: Output and Substitution Effects

When a firm employs two or more variable factors of production, the response of a firm to a change in factor prices is more complicated.

Table 7.1
Marginal Revenue Product of Labor in Sandwich Production (One Grill)

Total Labor Units (Daily Employees)	*Total Product (Sandwiches)*	*Marginal Product (Sandwiches)*	*Price of Output (per Sandwich)*	*Marginal Revenue Product Product*
0	0	—	$3	—
1	15	15	$3	$45
2	30	15	$3	$45
3	40	10	$3	$30
4	45	5	$3	$15
5	47	2	$3	$ 6
6	47	0	$3	0

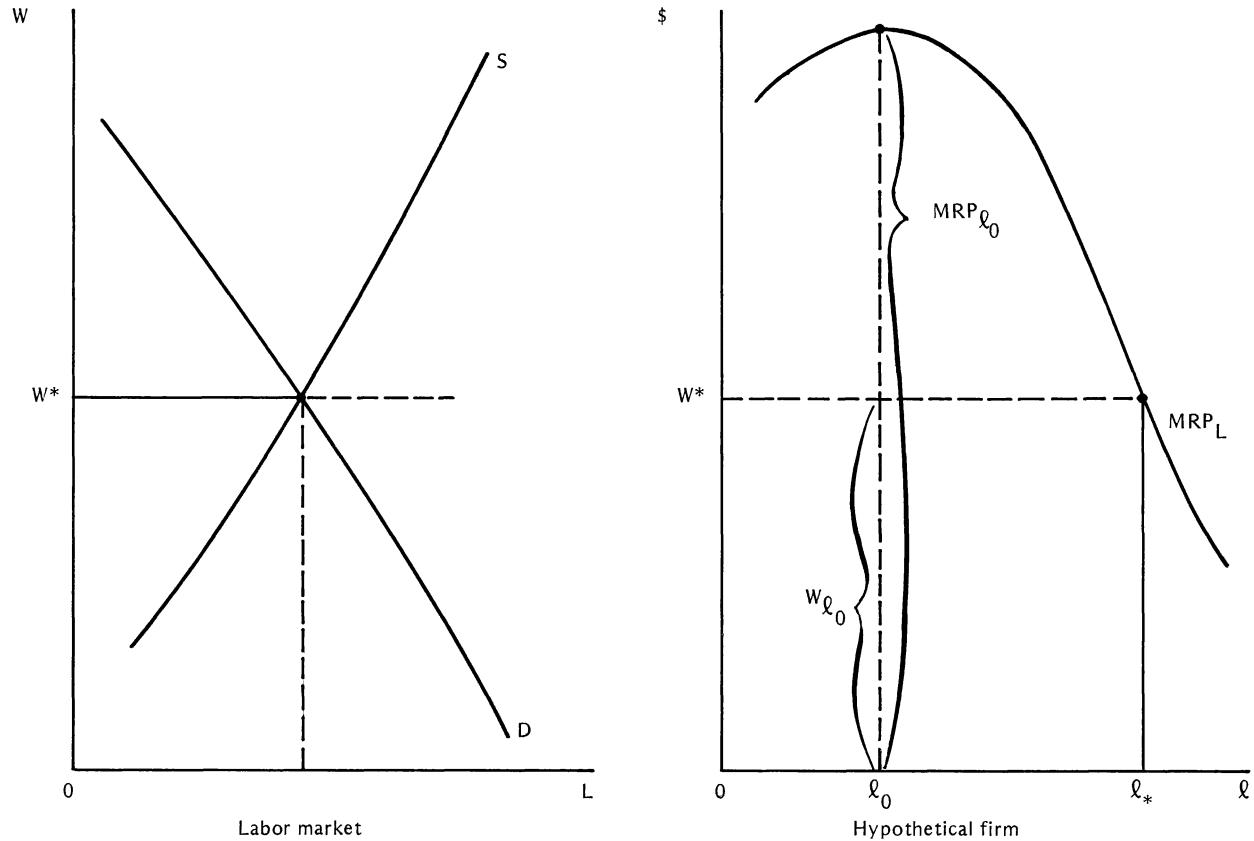


Figure 7.1. *Marginal revenue product and factor demand for a firm using one variable input (labor).*

In the following discussion we consider a firm that employs capital (K) and labor (L) and that faces factor prices P_k and P_l, but the logic applies to any two factors of production.

For now think of capital as physical capital (plant, equipment, and inventory) that is available at a given price per physical unit. In Chapter 9 we explore in more detail the idea of capital and the relationship between physical capital and the financial capital market. It will be argued there that through a set of complex institutional arrangements, firms that "demand" new physical capital in fact acquire it in a market where households provide the resources. Just as households supply and firms demand labor, households supply and firms demand capital.

Factors of production can be at the same time *complementary* and *substitutable.* That is, when two inputs are hired together, each enhances or complements the other. A new grill in the sandwich shop (capital) would raise the productivity of the sandwich makers (labor).

At the same time capital can be substituted for labor and vice versa. Some fast food chains have machines that practically make sandwiches by themselves. Recall from Chapter 5 that firms in most any industry you can think of have to choose among a variety of different production techniques; there are capital-intensive and labor-intensive methods of producing everything from automobiles to music. Changes in factor prices that can result from taxes may influence these choices.

Because of the complex technical relationship between inputs, firms may react in different ways to input price changes. First, when a factor's price rises relative to others, firms may find it advantageous to substitute other factors for it. High energy prices in the 1970s led firms and households to substitute capital (insulation, more efficient boilers, and so forth) for energy. High land prices lead to substitution of capital for land (tall buildings).

This effect is called the *factor substitution effect.* When the price of a factor rises, the substitution effect pushes firms in the direction of buying less; when the factor price falls relative to others, the substitution effect pushes firms to buy more of the factor.

At the same time, increases and decreases in factor prices will cause costs of production to rise and fall respectively. Cost increases may cause firms to cut back production; cost decreases may cause firms to increase production. Recall that the marginal cost curve is the supply curve of the profit maximizing firm; at a given price, a shift down (lower cost) will increase the profit maximizing level of output. When firms increase output, they tend to demand more of all factors of production. This is called the *output effect.*

Output and substitution effects together determine the overall change in demand for an input in response to a price change. A wage decline, for example, will influence labor demand in two ways. First, the firm

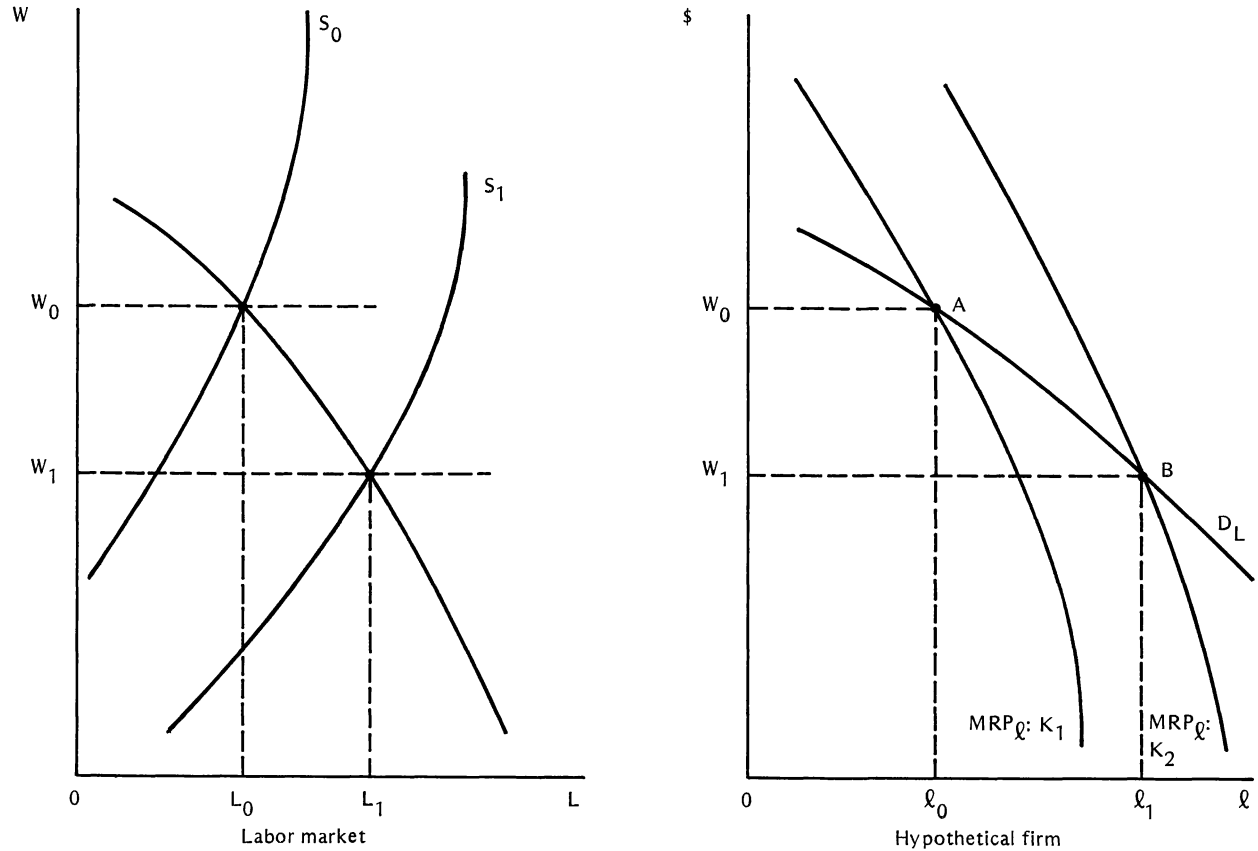


Figure 7.2. *Factor demand for a firm employing two variable factors of production (capital and labor).*

will tend to substitute labor for capital, increasing the demand for labor (factor substitution effect). Second, the lower cost of labor will reduce overall cost of production and lead to more output. The increase in output will also lead to more demand for labor (output effect).

When there are two variable factors of production, deriving the firm's demand schedules for inputs is more complicated. Figure 7.2 shows how a typical firm may react to a wage decline caused by an increase in labor supply. Initially, a firm facing a market wage w will hire as long as the marginal revenue product of labor is above w_o. Then l_o is the profit maximizing amount of labor to hire.

When the market wage falls to w_1, both output and substitution effects push firms toward hiring more labor. Recall, however, that *MRP* was defined as the revenue derived from hiring one more unit of labor *ceteris paribus*, holding other factors constant. Now, however, the lower cost might push firms to produce more output and hire more *capital* as well. With more capital, the marginal revenue product of labor shifts to the right as labor productivity is enhanced.

Using the example of the sandwich shop, lower wages and cost that reduce sandwich prices may mean a lot more business for our shop. If they buy a new grill, the *MRP* of sandwich makers will shift to the right. Table 7.2 and Figure 7.3 show the impact of lower sandwich prices and a "doubled capacity" (new capital) on the performance of workers.

The final labor demand curve will thus be the focus of all points like A and B in Figure 7.2. Notice that at every point of equilibrium,

$$W_o = MRP_l$$

Table 7.2
Marginal Revenue Product of Labor in Sandwich Production (Two Grills)

Total Labor Units & Daily Employees	*Total Product (Sandwiches)*	*Marginal Product (Sandwiches)*	*Price of Output (Per Sandwich)*	*Marginal Revenue Product*
0	0	0	$2.50	
1	15	15	2.50	$37.50
2	30	15	2.50	37.50
3	45	15	2.50	37.50
4	60	15	2.50	37.50
5	70	10	2.50	25.00
6	80	10	2.50	25.00
7	85	5	2.50	12.50

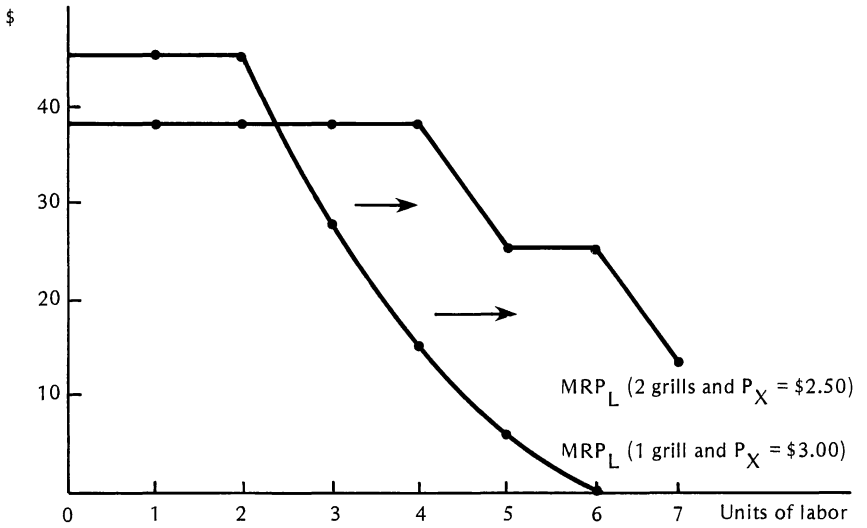


Figure 7.3. *Effect of increase in capital (grills) on labor productivity on sandwich production.*

Income Distribution and Marginal Productivity Theory

There are several important things to note about the view of input markets presented in this chapter. First, note that the compensation paid at equilibrium to each factor of production is linked to the worth of that factor's marginal product. The wage paid to labor, for example, depends on the value that society places on its "product" in output markets. Since compensation of factors in large measure determines the final distribution of income, this is often referred to as the marginal productivity theory of distribution.

Second, it should be noted that capital and labor are viewed as symmetric inputs. That is, both labor and capital are productive, and both are paid in accordance with the value that they add in production. You can produce more sandwiches with more sandwich makers or with a new grill.

This view of capital conflicts with the view implicit in other theories of distribution. For example, profit to capital in Marxian theory is not a reward for its usefulness in production; it is a wrongful appropriation of value actually created by labor.

Summary

Figure 7.4 shows the decisions that a competitive firm must make with respect to output supply and input demand using information on technologies available and on the market prices of inputs.

In diagram *A*, information on technologies and input prices is embodied in the marginal cost curve and firms will hire up to the point that output price (marginal revenue) is equal to marginal cost: $P_x^* = MC$.

In diagrams *B* and *C*, output price and technological information are embodied in the marginal revenue product curve, and firms will hire inputs as long as the factor's marginal revenue product exceeds its price:

$$P_l^* = MRP_l$$

and

$$P_k^* = MRP_k$$

It is important to remember that these are merely two perspectives on the same decisions. l^* and k^* must be fully consistent with Q_x^*. That is, using the least cost technology, l^* units of labor and k^* units of capital will produce Q^* units of output.

Review Questions and Exercises

1. Assume that a firm that produces TWITS can produce them with one of three processes used alone or in combination. The following table indicates the amounts of capital and labor required by each of the three processes to produce *one* TWIT.

	Units Labor	*Units Capital*
Process I	4	1
Process II	2	2
Process III	1	3

   a. Assuming that P_k = \$3 and P_l = \$1, which processes will be employed?

   b. Plot the three points on the firm's *TVC* curve corresponding to Q = 10, 30, and 50.

   c. At each of the three output levels, how much K and L will be demanded by the firms?

   d. Repeat b through c assuming that P_k = \$3 and P_l has risen to \$4.

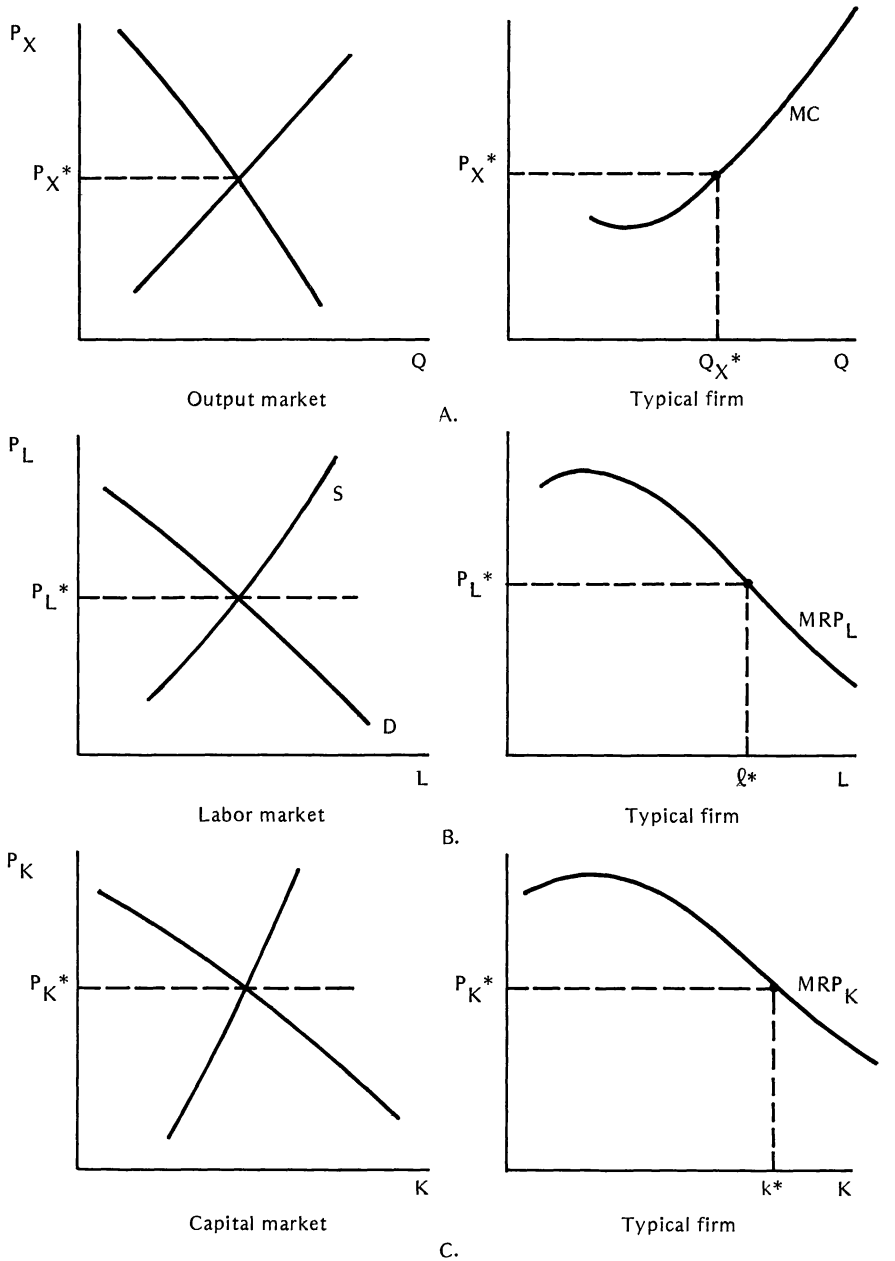


Figure 7.4. *Short-run decisions of a competitive firm.*

2. Assume that shoes are produced with one variable factor (labor). Using the following production function, estimate the amount of labor a competitive firm would hire assuming that labor is available for $35 per day and that shoes sell for $10 per pair.

Units of Labor (Days)	*Total Output (Pairs)*
1	5
2	9
3	12
4	14
5	15

3. Suppose that the government raised the payroll tax imposed on *employees* from the current level of 7.1 percent to 15 percent. Describe potential impacts on the demand for labor. Be sure to discuss output and substitution effects. Suppose the government raised the payroll tax imposed on *employers* from the current level of 7.13 percent to 15 percent. Describe potential effects on the demand for labor. Be sure to discuss output and substitution effects.

8

Competitive Pricing, Output, and Market Adjustment in the Long Run

Thus far, we have restricted our analysis to decisions made by firms in the short run. In the short run, there is some fixed factor of production that we usually think of as scale of operations, and firms cannot enter or leave the industry. There are two implications for cost: (1) Fixed costs must be borne in the short run even if the firm shuts down production, and (2) a firm attempting to increase output within its scale constraint faces diminishing returns and rising marginal cost of production. In the long run, however, all factors of production are variable, there are no fixed costs, and firms are free to move into and out of industries in response to economic profits and losses.

Economies and Diseconomies of Scale

Before turning to the behavior of individual firms, it is important to describe costs in the long run. The upward slope of the marginal cost curve and the U-shape of the average cost curve in the short run follow directly from the assumption of diminishing returns and fixed scale. Since scale of plant is infinitely variable in the long run, there is no assumption from which the shape of the long-run cost curves can be deduced.

The shape of a firm's long-run cost curve depends on how costs vary with scale of operations. It is quite possible that an increased scale or size will reduce costs, but increased size may also lead to inefficiency and waste.

It is important to distinguish between cost changes related to the size of individual firms and cost changes related to the size of the industry. Cost changes that result from the expansion of an individual *firm's* size are called *internal economies and diseconomies of scale.* The most common are technological in nature. If a firm expands and achieves lower average costs of production, it is said to have achieved internal economies of scale. Its production function is said to exhibit *increasing returns to scale.*

Technically, increasing returns means that if all inputs are multiplied by some number, A, output increases by a *larger* multiple. For example, if we doubled inputs and output went up 2.5 times, there would be increasing returns.

Figure 8.1 shows short-run and long-run average costs for a firm exhibiting economies of scale up to 150,000 units of output and *constant returns to scale* beyond that. Constant returns to scale implies that doubling inputs will double output, tripling inputs will triple output, and so forth.

The diagram shows three potential scales of operation. Each has associated with it a set of short-run curves because if that scale were built, the firm would be "locked in" to them in the short run. If the firm were to settle on scale I, it would not be realizing the significant cost advantages of producing on a larger scale. By roughly doubling its scale of operations to scale II, the firm can reduce average costs significantly.

Examples of internal economies of scale include moving to more capital-intensive automated production techniques. Public utilities operate much more efficiently on large scale, and large scale often means more "division" of labor and specialization.

Beyond 100,000 units, the firm in Figure 8.1 realizes only modest gains. Adding another 50 percent to plant capacity and moving to scale III does not materially affect the average cost of production.

If increasing the size or scale of a firm leads to *higher* costs, the firm is said to exhibit *decreasing* returns to scale or internal diseconomies of scale. One can imagine coordination problems or bureaucratic inefficiency. Technically, decreasing returns implies that doubling all inputs leads to less than twice as much output, and so forth.

Often costs facing individual firms can rise or fall when the *industry* expands. Cost savings resulting from expansion of an industry are called *external economies of scale*; such industries are called decreasing cost industries. If a new and expanding industry demands skilled labor, initially in short supply, labor costs may actually fall as people are trained in larger numbers over time. This has nothing to do with the size of individual firms.

If costs rise as an industry expands, there are external diseconomies of scale; those industries are called increasing cost industries. Rapid

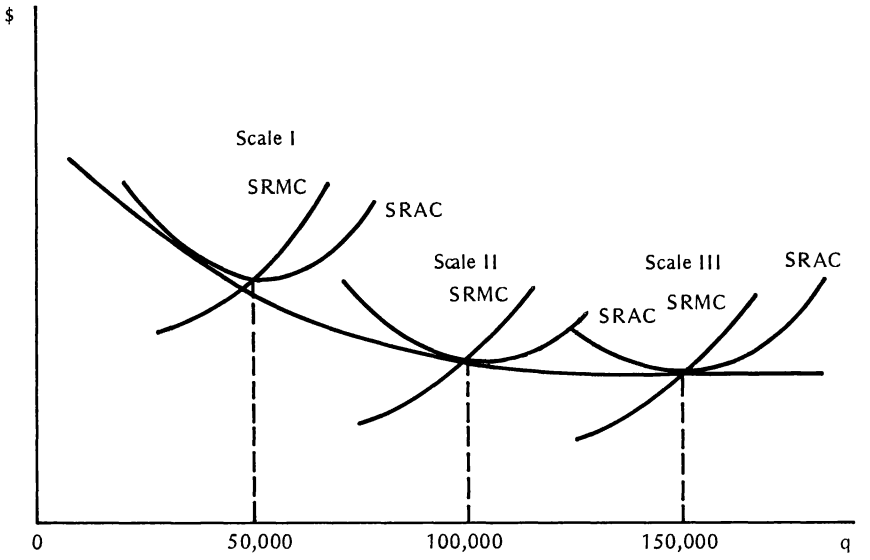


Figure 8.1. *A firm exhibiting internal economies of scale.*

expansion of the construction industry, for example, can drive up the price of lumber for all firms even though the firms themselves are quite small.

Expansion of Firms and Industries in Response to Economic Profits

Figure 8.2 shows an industry in which individual firms are earning economic profits. Total revenues exceed total costs, and firms are earning profits over and above a normal rate of return. In the long run, there is an incentive for two types of response: (1) Existing firms will probably find it in their best interests to expand; and (2) new firms will begin to enter the industry to reap the higher-than-normal profits available.

With price in excess of cost, existing firms will expand even if there are no economies of scale. Figure 8.3 shows such a firm. Assume that the firm begins producing in a plant of scale I, facing a market price of P_o. In the short run, a typical firm will produce at q_o and earn economic profits shown in the indicated shaded area. Assuming that the market price remains fixed at P_o, increasing scale to scale II in the long run will more than double economic profits.

The assumption that price will remain at P_o, however, is a bad one.

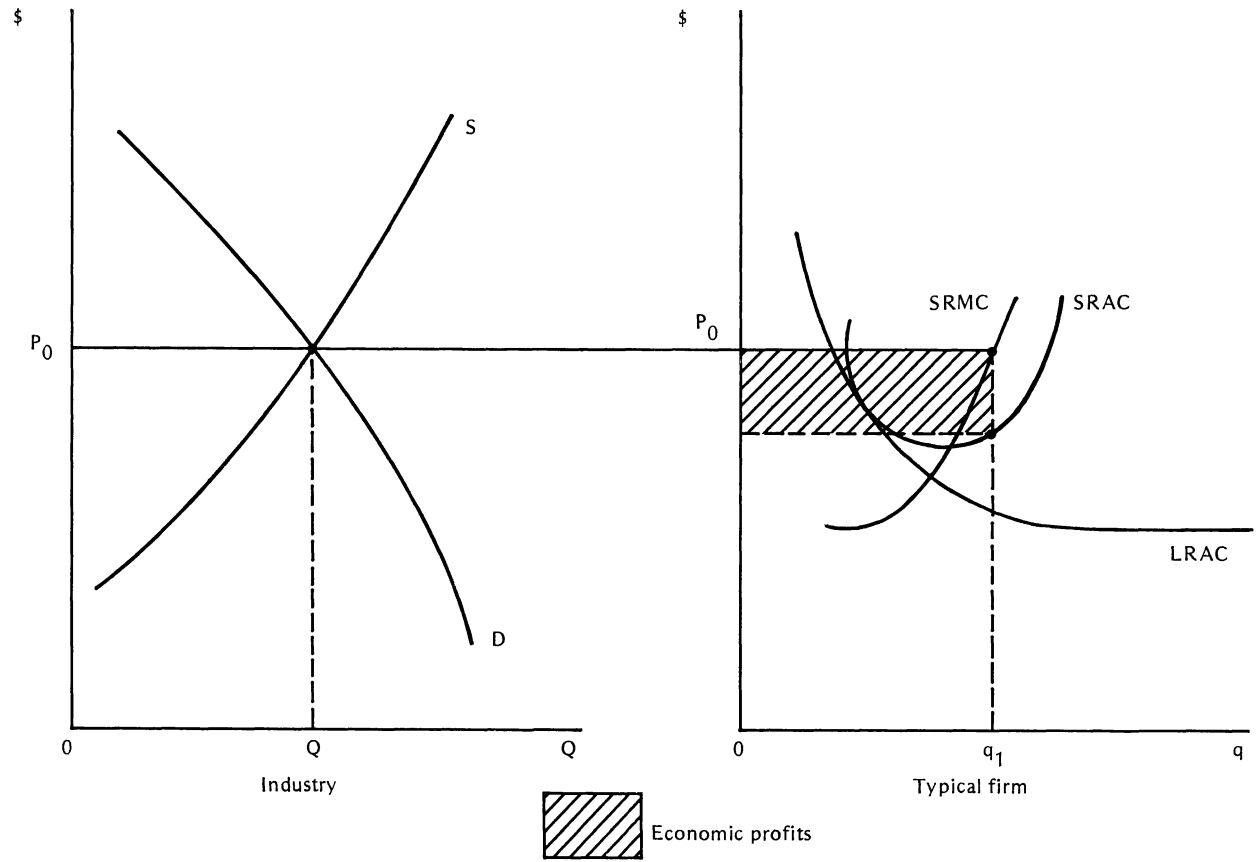


Figure 8.2. *Firm exhibiting increasing returns to scale earning economic profits.*

Recall that the market supply curve is the horizontal sum of the marginal cost curves (above-average variable cost) of all the firms in the industry. With economic profits and competitive free entry, industry supply will shift to the right, driving down market price. First, new firms will enter and more will be supplied at each potential price. Geometrically it simply means there are more marginal cost curves to add up. Second, remember that existing firms have an incentive to expand their scales of operation. Notice that our hypothetical firm has shifted its short-run cost curves substantially to the right. The expansion of existing firms, then, also contributes to the outward shift of the industry supply curve.

As long as economic profits exist, there will be an incentive for capital to flow into the industry. Thus, the industry supply curve will continue to shift out, driving down the price of output until all economic profits have been eliminated. In Figure 8.3, this occurs at price P^*. At long-run competitive equilibrium,

$$P = SRMC = SRAC = LRMC = LRAC$$

Economic profits are zero.

Critics of orthodox economics point out that the final structure of an industry cannot be determined from this scenario. While we have been assuming a "competitive structure" (that is, a large number of firms each small relative to the market), there is nothing to suggest that this will be the result. If existing firms expand faster than new firms enter in response to profits, concentration may result. We will discuss problems associated with more concentrated industries in Chapters 11 and 12 below.

Contraction of Industries Suffering Short-Run Losses

Figure 8.4 shows an industry in which existing firms producing with constant returns to scale are suffering short-run losses. In the short run, since the initial price, P_o, is above average variable costs, firms will continue to produce generating a profit on operation (revenues in excess of variable cost). The profit on operation can be used to offset fixed costs and reduce losses.

In the long run, however, firms have incentive to reduce scale and exit the industry. New capital will not be invested, and some firms will liquidate and go out of business. As they exit, the supply curve will shift to the left creating a temporary shortage and driving up price. While it does not show on the diagram, existing firms may also reduce scale shifting their short-run curves to the left and accelerating the downward shift of the industry supply curve.

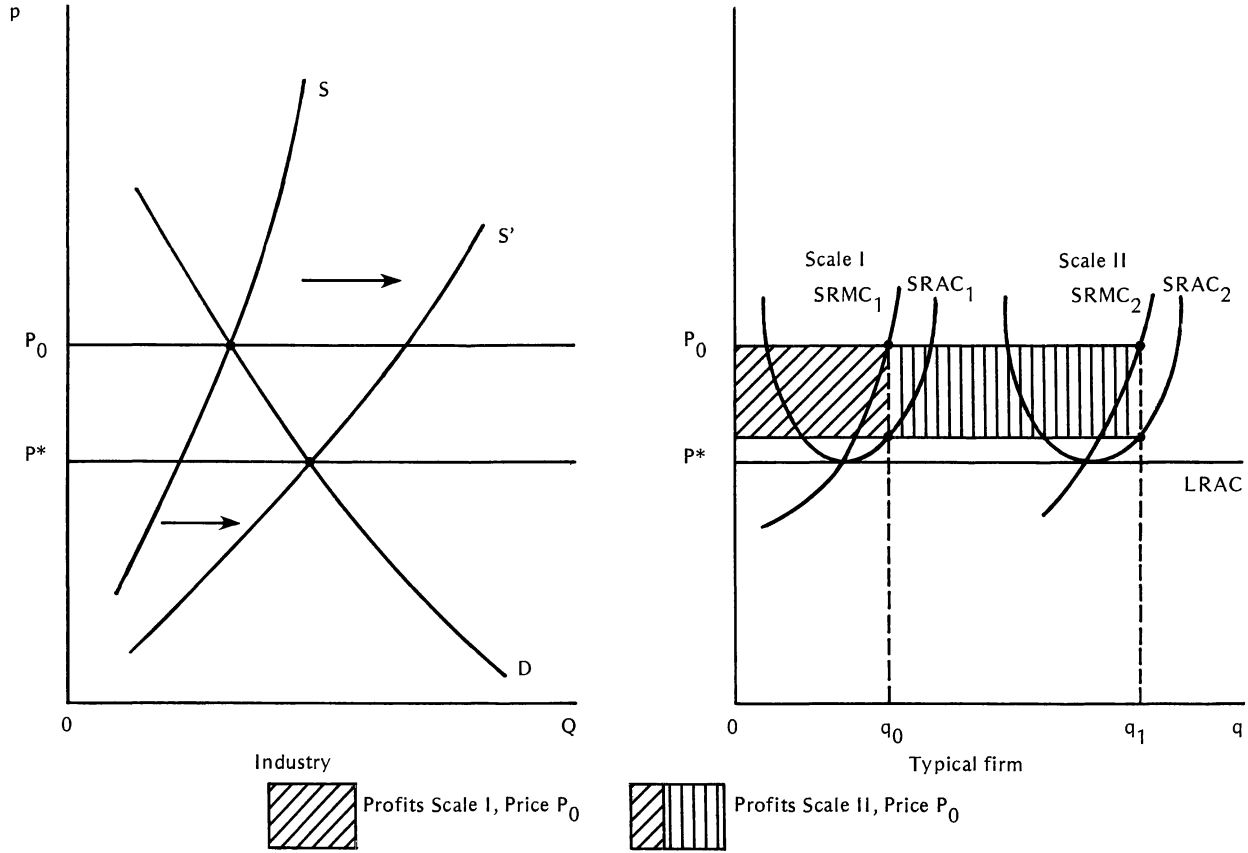


Figure 8.3. *Long-run expansion and entry in an industry with constant returns to scale.*

Supply will continue to decline, driving up prices until losses have been eliminated. This will occur when

$$P^* = SRMC = SRAC = LRMC = LRAC$$

The Market System as an Allocative Mechanism

At this point we have, in a sense, closed the circle. We have examined briefly the process of economic decision making in households and firms and have discussed the basic operation of competitive input and output markets. We have a description of a simple market system that provides some answers to the questions we asked at the outset. The argument at the beginning was that every society has some mechanism or system that determines the *allocation* of resources among productive agents, the *mix* of output, and the *distribution* of that output among its citizens. Stated another way, every society must decide *what* will be produced, *how* it will be produced, and *for whom*.

We also argued at the outset that in every society markets exist. The first eight chapters of this book have built a model of the simple market system under the assumption of perfect competition. They described forces that operate both in simple and more complex systems and began to explain how markets might be expected to react to tax policy measures.

To review, consider the response of a system to a change in consumer preferences (see Figure 8.5). First, we know that the effective demand that a household is able to exercise is constrained by income and wealth. The distribution of income and wealth is determined by two things: (1) the factors of production that households own (land, capital, and skills) and (2) the prices that they command in the market. The initial distribution of factors is determined outside the system being analyzed. Wealth may be inherited, stolen, or simply saved out of previous income.

Suppose households acquire a taste for video games. Those with the acquired taste and sufficient income go to the market and exercise effective demand. Increased demand causes prices to rise, and firms will earn economic profits. Higher prices cause output to increase, but firms are constrained by their current capacities (scales of plant) in the short run.

In the long run, existing firms will expand, and new firms will enter the industry. Demand for specific inputs will increase. For example, programmers who are trained in complex graphics are needed to produce video games. Increased output price and new firms will shift input demand curves (*MRP* schedules) up. As a result, the prices of

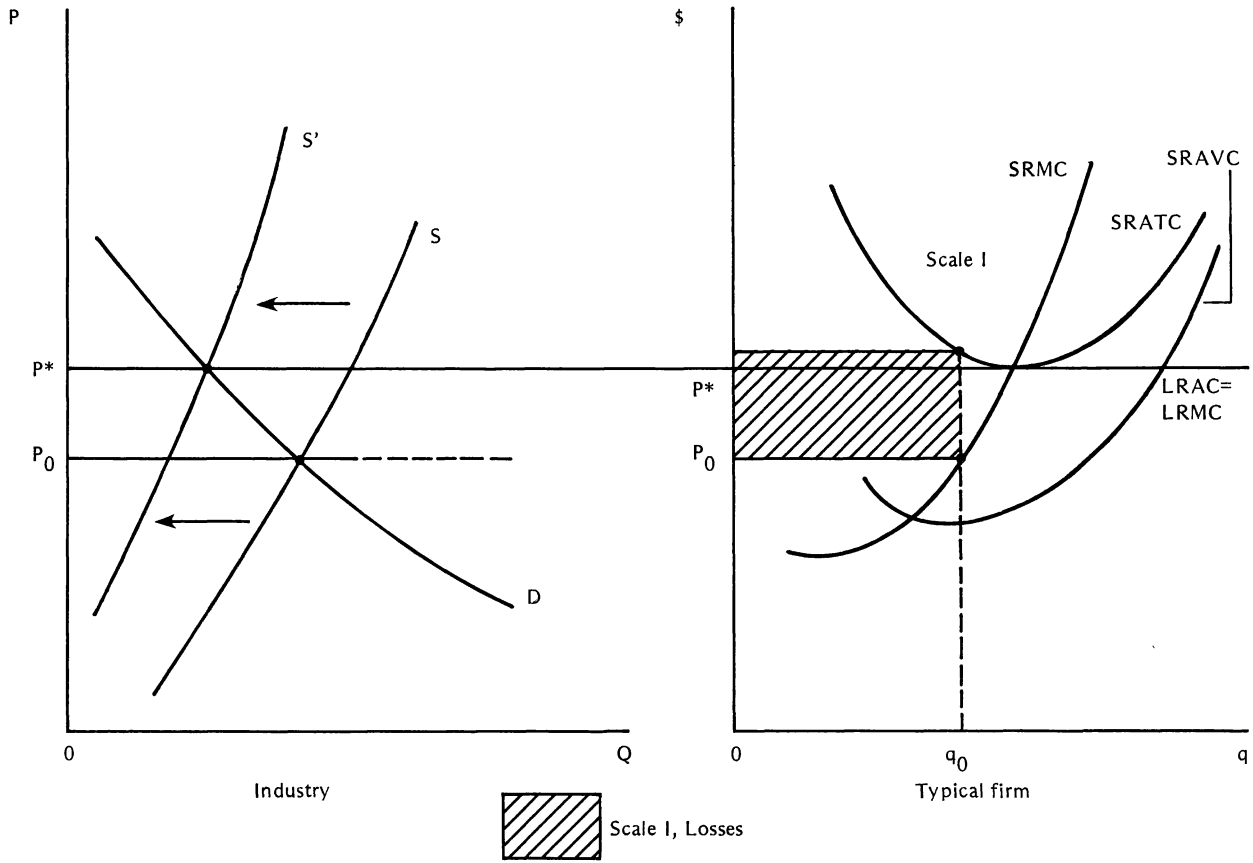


Figure 8.4. *Long-run contraction and exit in an industry with constant returns to scale.*

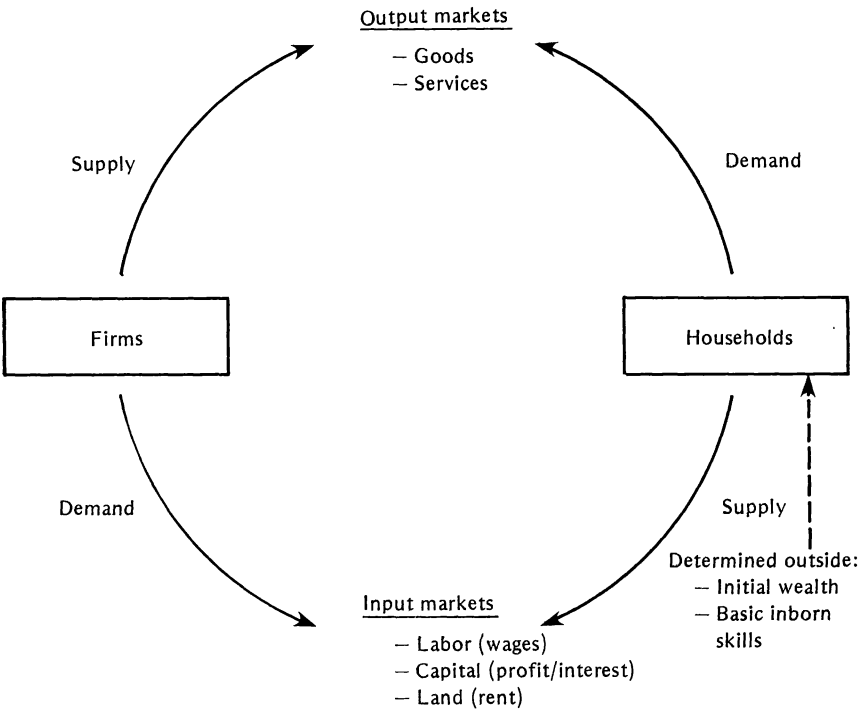


Figure 8.5. *The market allocation mechanism.*

factors needed will rise, giving people the incentive to acquire those skills. Expanding firms will also be able to bid more for good locations, and land may be reallocated. Ultimately, resources have moved into the production of video games, production proceeds with the most efficient technology, and firms enter until the final market price is driven back to cost.

Taxation and Interindustry Resource Flows

Taxation and tax policy can have an effect that is similar to the one described for a shift of preferences. Suppose, for example, that we imposed a tax of 50 percent on the profits of one large sector of the economy. One might think, for example, of a tax on luxuries or on the corporate form of business organization.

If we begin with all firms in long-run competitive equilibrium, the tax reduces the profits in the taxed sector to below-normal levels. With economic losses (below-normal profits) firms in the taxed sector begin to contract or exit. Prices in the taxed sector rise, and consumers of

the taxed sector's goods bear some of the burden of the tax.

As capital leaves the taxed sector, it flows into the nontaxed sector, since at least normal profits are available there. New firms enter, and existing firms expand. Prices in the nontaxed sector fall, and profits are driven down. The process stops when the after-tax profit rate in the taxed sector and the no-tax profit rate in the nontaxed sector are equal.

To the extent that the taxed and untaxed sectors use different inputs, the relative prices of inputs may adjust and income may be redistributed.

It is clear that the long-run impact of taxes on the economy may be quite complex. Chapter 10 explores the topic of tax incidence in more detail.

Review Questions and Exercises

1. Suppose that firms in industry X had production functions that exhibited increasing returns to scale up to 50,000 units of output but decreasing returns to scale above that level. Sketch a typical firm's long-run average variable cost function. If each firm began at a small scale, but earned economic profits, at what scale would they ultimately settle?
2. Suppose that there were only two sectors of an economy, one of which used energy extensively. Suppose that the government, in its desire to achieve independence from foreign sources, placed a large tax on energy use. Describe in detail the long-run adjustment (including possible factor substitution) likely to follow imposition of such a tax.
3. Consider an industry that exhibits *external* diseconomies of scale (an increasing-cost industry). Suppose that demand for its product were to increase significantly. Describe in detail the adjustments likely to follow. Where would market price ultimately settle? Use diagrams in your answer.

9

Capital and Land Markets

In Chapter 8 we discussed the expansion and contraction of firms and industries in response to economic profits and losses. When firms expand, they put physical capital in place in the form of new plant and equipment, and they hire additional variable inputs. We have been saying loosely that firms demand and that households supply inputs: land, labor, and capital. Transactions between households and firms in labor and land markets are fairly direct and easy to conceptualize. Households offer labor services directly to firms in exchange for a wage; landlords or landowners either sell or lease land directly to firms in exchange for rent or an agreed on price.

This connection between households and firms in the capital market is much less direct and difficult to understand. Firms purchase or demand capital assets but not directly from households; households supply the financial resources necessary. The mechanism involves a complex set of institutions (the stock market, banks, venture capital funds, brokerage houses, and so forth), collectively referred to as the *financial capital market.*

The chapter describes and begins to analyze the basic functioning of the capital market. We first describe more precisely the nature of *capital, investment,* and *depreciation* and then turn to the mechanics of individual investment decisions. This material is of tremendous import to tax policy makers.

Capital, Investment, and Depreciation

There is no more important concept in all of economics than *capital.* Capital goods (1) yield valuable services over time; (2) are produced

by the economic system itself, and (3) are used as inputs to produce other goods and services. While one generally thinks of physical capital employed by business firms, such as plant and equipment, there are intangible forms of capital as well. A firm that invests in advertising to establish a "brand name" is producing intangible capital called goodwill: Presumably that goodwill will provide benefits to the firm into the future.

Households may also produce or use both tangible and intangible capital. If I build a vacation home for myself, it is capital that will produce housing services for me in the future. If I invest in an education, I may be acquiring human capital that will provide me with valuable services through time.

The most important dimension of capital is the time dimension. The decision to put new capital in place must by its nature be made with an eye to the future. The value of an investment in capital is only as great as the value of the services it will produce over time; in a very real sense, the capital is the physical embodiment of those future services.

Since capital comes in many forms, it is impossible to measure directly in physical terms. As an indirect measure, we generally use its present market value. The measure of a firm's *capital stock* is thus the current value of its plant, equipment, inventory, and intangible assets. This allows us to add machines to goodwill and to make cross-firm comparisons. It does have the disadvantage of deemphasizing the heterogeneous nature of the capital stock.

Capital is measured as a stock. That is, like the volume of water in a tub, it is measured at a point in time. Stop the clock and measure: The capital stock of the *XYZ* corporation on July 15 is $3,425,000. Although it is measured in money or value terms, it is very important to think of the actual stock itself.

Stocks of capital are affected over time by two flows: *investment* and *depreciation*. When a firm produces or puts in place new capital, such as a new piece of equipment, it has "invested." Investment is a flow; it has a time dimension. We speak of investment per period (per month or per year, and so forth).

The act of "investing" should be distinguished from the act of buying a share of stock or a bond. Although the term is often used to refer to the latter, and despite the fact that real investments are often made with the proceeds of a stock sale, the correct use of the term *investment* refers to the creation of capital.

Depreciation refers to the erosion or deterioration of a capital asset over time. Its correct measure is the decline in economic value of the asset. Depreciation is also a flow; it has a time dimension. To pursue the useful analogy with the volume of water in a tub, investment is

like the flow of additional water from the tap per hour, and depreciation is like the amount of evaporation per hour.

The Investment Decision: Demand for Capital

While the decision process discussed here will be described from the standpoint of a business firm, it is important to remember that investment decisions are made by households and governments as well. The basic nature of the decision is the same for all. Capital assets yield valuable future services, and any decision to purchase a capital asset must involve an analysis of them. A firm must evaluate the future profits or revenues that a new plant will generate; a household must evaluate the value of future services that a new automobile will provide; governments must evaluate the value of benefits that society will derive from a new bridge or a war memorial.

First of all, any such evaluation involves forecasting. Decision makers must have expectations about the future in which the asset will exist. A new plant will be worth a lot if the market for a firm's product grows and its price remains high, but it may be worth little if the market declines or if the price of its product falls.

An official of the General Electric Corporation once described an investment decision facing his firm. Economic forecasters had provided the company with ten-year predictions of "new housing construction" that ranged from a low of 400,000 units per year to a high of 4 million units per year. Since General Electric sells millions of household appliances to contractors building new units, the forecast was critical. If the low number was correct, it meant that GE would close several of its larger plants. If the high number were correct, it meant literally billions in new plant and equipment investment over the next ten years. Households must make similar judgments about the future. A home on a wilderness lake built for purposes of getting away from overcrowded population centers will be worth much less if the wilderness area becomes overdeveloped and fifty motor boats traverse the lake daily.

For a firm, the second step involves calculating the cost of the investment project. We will assume that business firms have access to financial markets and that at any moment there is a market clearing *interest rate*. In simplest terms, the interest rate is the price that borrowers pay to lenders for the use of their money. Although the actual contract interest rates on a given loan differ with length and perceived risk of loans, there is generally a prevailing range, and we speak of "the" interest rate.

If we assume that firms can either borrow or lend at the prevailing interest rate, then interest can be thought of as the *cost of capital*. If a

firm finances an investment project by borrowing over the life of the project, the cost that is borne is simply the interest paid on the loan. Similarly, however, if the firm were to finance the project out of on-hand cash, it would *forgo* the interest that the cash would earn if it had been loaned out (an opportunity cost). Thus, the current interest rate is an approximation of the cost of capital.

Once expectations have been formed, firms must quantify them. One way is to calculate an *expected rate of return* on the investment project. (The actual method of calculation is discussed below.) If I invest in a new plant that costs $1 million to put in place and I *expect* it to raise my profits by $200,000 per year from now on, the *expected rate of return* is 20 percent.

Table 9.1 presents a menu of investment choices facing a hypothetical firm. Expected rates of return are based on forecasts of future profits attributable to the investments. Figure 9.1 is a graph of the total amount of the investment in millions of dollars that the firm would undertake at various market interest rates. At an interest rate of 25 percent, no projects would be undertaken. The cost of borrowing (or the forgone interest) would exceed the return on all the projects. At a market interest rate of 16 percent, however, the first three projects would be done at a total cost of $2.9 million. Similarly, at 10 percent the first five would be undertaken at a cost of $4.1 million.

The investment schedule is in essence the firm's demand for capital as a function of the market interest rate. The horizontal sum of all the individual firm's investment schedules is the demand for capital in the economy as a whole (see Figure 9.2). Keynes called this curve the *marginal efficiency of investment* function. In a sense, it is a ranking of all the investment projects in the economy in order of expected yield.

The most important single thing to remember about this curve is that its shape and position depend critically on the *expectations* of those

Table 9.1
Hypothetical Firm's Possible Investment Projects Ranked by Expected Profitability

Project Number	*Name*	*Cost*	*Expected Rate of Return*
1	New plant	$1,500,000	24%
2	Fleet of trucks	$400,000	21%
3	Modernize old plant	$1,000,000	17%
4	Advertising campaign	$500,000	14%
5	New computer	$700,000	12%
6	Office equipment	$300,000	9%
7	Second new plant	$1,000,000	3%

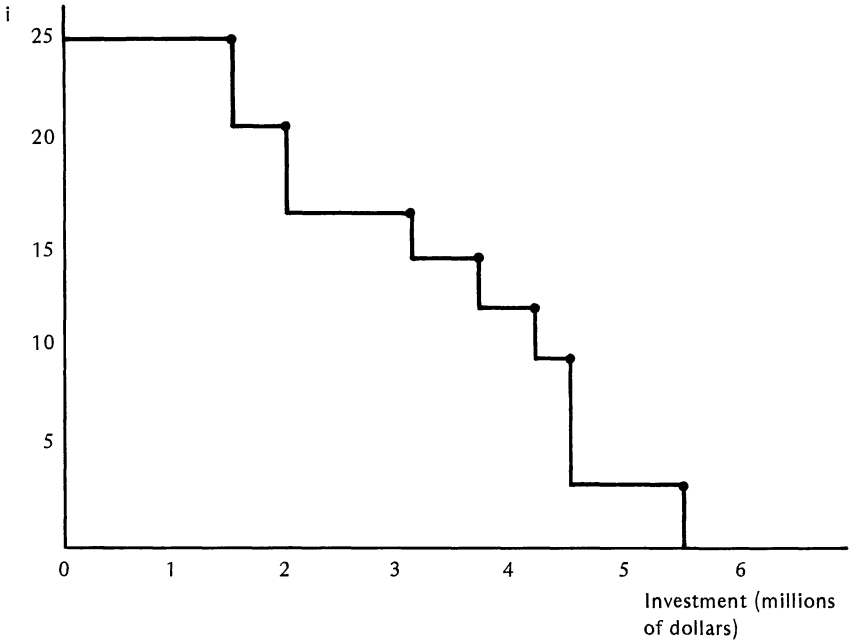


Figure 9.1. *Investment schedule for a hypothetical firm.*

making investment decisions. Those expectations depend on many things, and in all probability they are volatile and may change frequently. Clearly, upswings and downswings (recessions) in the economy as a whole will have an influence. Uncertainty always works against expectations. Keynes said that investment depends on "animal spirits of entrepreneurs." So while lower interest rates tend to stimulate investment and higher interest rates slow it, other hard-to-measure and hard-to-predict factors also play an important role.

Discounting and Present Value

One way to evaluate a potential investment project is to calculate the expected rate of return and then to compare it with the market interest rate. This section of the chapter discusses a nearly equivalent method of analysis.

The basic problem, of course, is how to evaluate streams of value or income that will materialize only in the future. Consider, for example, the expected flow of profit from an investment project in Table 9.1a. If such a project cost $1,200 to put in place, would the firm undertake it? At first glance you might answer yes, since the total flow of profit is $1,600. But that flow of profit comes only after a period

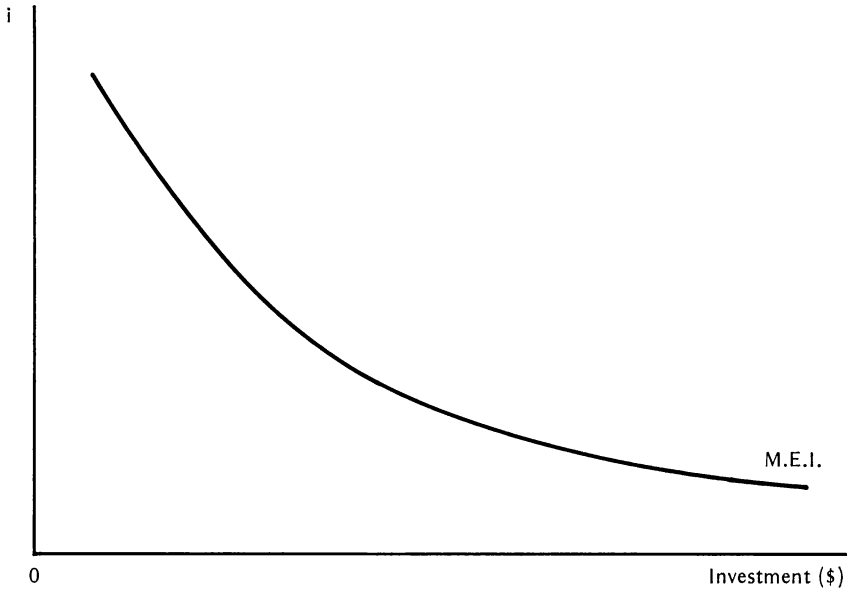


Figure 9.2. *Investment demand or marginal efficiency of investment.*

of years: the $1,200 could be put safely into a money market account where it would earn interest, and I would end up with more than I have now. Intuition should tell you that the desirability of this project depends on the interest rate available in the market.

One way of thinking about the availability of interest is that it allows us to sell and purchase future dollars. That is, I can buy a contract for a dollar to be delivered in one year by simply loaning out some amount of money today. An interesting question is: How much do I need to pay today (loan out) in order to buy a contract for delivery of one dollar in one year? The answer is quite simple: I must pay X

Table 9.1a

End of Year . . .	*$*
1	100
2	100
3	400
4	500
5	500
All later years	0

such that $X + iX$ is equal to \$1.00 where i is the interest rate and iX is the interest earned in a year. Solving for X algebraically,

$$X + iX = 1 \rightarrow X = 1/(1 + i)$$

X is said to be the *present value* of \$1 one year from now. It is actually the current market price of \$1 to be delivered in a year.

What then is the present value of \$100 in two years? It is X such that X plus the interest it would earn, compounded for *two* years is equal to \$100. After one year, we would have X plus iX or $X \cdot (1 + i)$. After two years we would have $X \cdot (1 + i) + i\,[X(1 + i)]$ or $X(1 + i)(1 + i) = X(1 + i)^2$. Again, solving algebraically for X,

$$X = \$100/(1 + i)^2$$

If the market interest rate were 10 percent or .10, $X = \$100/(1.1) = \$82.64$. If I put \$82.64 in an account today at 10 percent, I would after two years have exactly \$100 in that account. Thus, the present value of \$100 two years from now is \$82.64.

In general, the present value of R dollars t years from now is

$$\frac{R}{(1 + i)^t}$$

Table 9.2 calculates the present value of the income stream above at an interest rate of 10 percent. The total present value turns out to be \$1,126. This says that if I simply go to the financial market today, I can buy a contract for \$100 one year from now and a contract for

Table 9.2
Calculation of Total Present Value of a Hypothetical Investment Project (assuming i = 10%)

End of Year. . .	*\$*	÷	=	*Present Value*
1	100	(1.1)		90.91
2	100	$(1.1)^2$		82.64
3	400	$(1.1)^3$		300.53
4	500	$(1.1)^4$		341.51
5	500	$(1.1)^5$		310.46
-	0			
Total Present Value				\$1,126.05

$100 two years from now and a contract for $400 three years from now, and so forth; all for the low price of $1,126! Why then would I pay $1,200 to do this investment project? Answer: I wouldn't.

But suppose I later discover that interest rates have fallen to 5 percent. If that's so, contracts for future dollars are more expensive. Since I would earn less interest, I would have to pay more today. Table 9.3 recalculates the present value of the same stream at a market interest rate of 5 percent. Total present value rises to $1,334, and the project, at a cost of $1,200, seems like a good deal.

If project cost exceeds *present value* at current rates, it should not be undertaken; if present value exceeds present project cost, it should be undertaken.

Asset Values, Present Value, and Future Income Streams

It was argued above that the value of investment in capital is only as great as the value of the services it will produce over time. We now can say more formally that the *market price* of an asset is generally driven to the *present discounted value* of the stream of earning or services that the asset is expected to produce over time.

It is easiest to see how the market does this by considering a *bond.* A bond generally represents a fixed, guaranteed stream of earnings to be paid over time. For example, an 8 percent, 10-year, $1,000 bond entitles the owner to $80 a year for ten years and $1,000 back at the end of ten years no matter what happens to the market interest rate.

Table 9.3
Calculation of Total Present Value of a Hypothetical Investment Project (assuming i = .05)

End of Year . . .	*$*	÷	=	*Present Value*
1	100	$(1.05)^1$		95.24
2	100	$(1.05)^2$		90.70
3	400	$(1.05)^3$		345.54
4	500	$(1.05)^4$		411.35
5	500	$(1.05)^5$		391.76
Total Present Value				$1,334.59

If I bought such a bond, the company or government that issued it did so to borrow my money, and there is no obligation to pay it back for ten years. There is however, a market for previously issued bonds; they clearly have a value because they represent a claim on future dollars. Their value derives directly from the future stream.

Suppose that I desired to sell my bond one year after I bought it but that interest rates are now 12 percent. How much would a "rational" person be willing to pay for my bond? She would certainly not pay $1,000 because now $1,000 would buy her $120 per year and $1,000 back, which is significantly more than $80 per year.

Table 9.4 shows that she could, with interest rates at 12 percent, buy the *exact same income stream* for $786.87. If she paid more than that today for my bond, she is in fact accepting a *yield* of less than 12 percent; if I offered it to her for less, she would earn *more* than a 12 percent yield. Be sure you understand why.

Millions participate in the capital market. There are literally thousands of money managers, pension funds, mutual funds, brokerage houses, and so forth, whose job is to find a "good" deal for households who have accumulated wealth. Everyone wants the best possible yield, and most are willing to pay to find it. Managers are continually scanning the financial horizon for profitable investments, and that process does two things: (1) It channels capital into its most productive use, and (2) it drives the value of marketable assets (physical and financial) to the *present value* of the future stream of earnings that the asset is likely to produce. I am not likely to find someone who will overpay me for my bond.

Table 9.4
Present Value of an 8 Percent, $1000 Bond When Interest Rates Rise to 12 Percent

End of Year...	*$*	÷	=	*Present Value*
1	80	(1.12)		71.43
2	80	$(1.12)^2$		63.78
3	80	$(1.12)^3$		56.94
4	80	$(1.12)^4$		50.84
5	80	$(1.12)^5$		45.39
6	80	$(1.12)^6$		40.53
7	80	$(1.12)^7$		36.19
8	80	$(1.12)^8$		32.31
9	1080	$(1.12)^9$		389.46
Total Present Value				$786.87

Example: Rental Property

Property tax assessors continuously make judgments about asset values. Often the asset in question has never been through the market. If, however, the asset is generating income, its value can be inferred from the stream of income associated with it.

Consider a large house in Cambridge, Massachusetts. Assume that the house is in good shape and that it has ten rooms that can be rented to students at $200 each per month. That would, with full occupancy, generate $2,000 per month or $24,000 per year in revenues to the owner.

Assume further that maintenance, repair and property taxes costs a total of $6,000 per year. The maintenance and repair are sufficient to maintain the value of the property indefinitely. How much is the house worth if these flows were to continue unchanged for a long period of time?

To the owner, the profit stream is $24,000 minus $6,000 or $18,000 per year. If interest rates were 10 percent and there was no risk associated with the investment, a rational investor would pay $180,000. That is because $18,000 is 10 percent of $180,000, and to pay more would imply a return of less than 10 percent; to pay less would imply a return of more than 10 percent. In fact, the present value of $18,000 per year for many years using a 10 percent discount rate converges to $180,000! (Can you prove that?)

Land and Land Value

Since land in general and each parcel in particular is fixed in supply, price is *demand determined.* The return to any factor of production in fixed supply is called a *pure rent.* (See Figure 9.3.)

Any given parcel of land may have a number of alternative uses. As with any other factor it will presumably in a market setting go to the potential user who is willing to pay the most for the site. The value of land to a potential user may derive from the characteristics of the land itself or from its location. For example, it may be very fertile and thus of great value in agriculture, or it may be at an accessible location and thus of great value as a commercial or residential site.

Land use patterns clearly change over time as economies grow and develop. Urbanization, population growth, and rising productivity in agriculture inevitably lead to a shift of land out of agriculture and into industrial and residential use.

The process of land allocation in the absence of regulation is no different than the process of labor or capital allocation except that supply

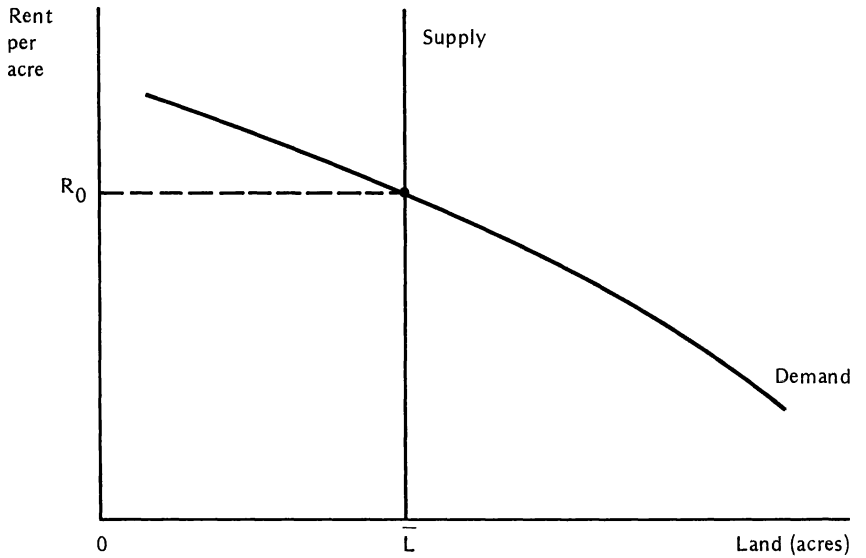


Figure 9.3. *The determination of rent.*

is perfectly inelastic. Any potential user will bid for it as long as the bid is below the marginal revenue product of the land.

As a simple example, consider two potential uses for a plot of land. Use *A* will require a $100,000 capital investment and yield a profit of $20,000 per year after maintenance, depreciation, and taxes but before rent is paid. Use *B* will require a $150,000 capital investment and will yield a $50,000 flow of annual profits. If we assume that the interest rate (or the normal rate of return to investment in the sectors involved) is 10 percent, then we can calculate the maximum bid that the land owner should be able to extract from each potential user.

In use *A* $10,000 of the $20,000 annual profit is required for the compensation of capital. Since a capital investment of $100,000 is required, 10 percent ($10,000) is by definition just sufficient to interest investors in the sector. Alternatively, the capital might be purchased with the proceeds from a loan; in that case the simple interest on the loan would be $10,000.

If the location enables the owner to earn the full $20,000, the landowner should be able to demand up to $10,000 per month for access to the location for use *A*.

In use *B* capital must receive $15,000 (10 percent of $150,000), but profits are $50,000 leaving $35,000 for land. Thus land use *B* will generate significantly more rent and presumably the land will end up there.

Another way to look at the same problem is to ask, If interest rates are 10 percent, how much would an investor be willing to pay in total

(land and capital) to receive a flow of $20,000 in profits? Since $20,000 is 10 percent of $200,000, an investor should be willing to pay up to that amount. Any amount less than $200,000 will mean that the investor is earning more than 10 percent. Since the capital costs $100,000, the full value of the plot of land in use *A* is $100,000 or $200,000 total minus $100,000 for capital. Notice that $100,000 is exactly equal to what a land buyer would be forced to pay to earn a rental stream of $10,000 per year. Stated alternatively, $100,000 is the present discounted value of the $10,000 annual rental stream.

Using the same logic in use *B* we arrive at a total value of $500,000, since the gross yield is $50,000. Since $150,000 is required for the capital, land can claim a total of $350,000. Again this will clearly be the highest bid; it is the present value of the $35,000 rental stream that would flow from use *B*.

Finally, it should be noted that since land cannot be physically moved, the value of any one parcel will depend to a large extent on the uses to which adjoining parcels are put. A smoking factory will likely reduce the value of adjoining land, while a new highway that increases accessibility may enhance it.

If no restrictions are placed on land use, individual users may have no incentive to consider these external effects (see Chapter 11). If land use *X* is $100 per month more profitable than land use *Y*, but it creates external costs of $200 per month, misallocation may occur if those costs are not considered.

Land use planning and zoning restrictions are used in almost all jurisdictions to ensure that such misallocations are avoided.

Review Questions and Exercises

1. Calculate the present value of the following streams at interest rates of 8 percent and again at 10 percent:

End of Year...	*a*	*b*	*c*	*d*	*e*
1	80	80	100	100	500
2	80	80	100	100	300
3	80	80	1,100	100	400
4	80	80	0	100	300
5	1,080	80	0	100	
6	0	80	0	1,100	
7	0	1,080	0	0	

2. If investment e in problem 1 were a machine that cost $1,235, would you purchase it assuming interest rates were 8 percent? 10 percent? Explain your answer.
3. The owner of a piece of land has determined that it has two potential uses: (1) He could build an apartment building on it for $150,000 containing ten apartments that would rent for $200 each per *month*; it would cost a total of $4,000 *annually* to maintain and operate. (2) He could build an office building for $130,000 with ten office suites that would rent for $150 per *month* each. It would cost $3,000 *annually* to maintain and operate the offices. Assume that interest rates are 10 percent.

   The current zoning law prohibits using the land for residential purposes. Thus, the apartment building is not an option. If an unscrupulous member of the zoning board agreed to change the zoning in exchange for cash, how much would the owner of the building be willing to pay, assuming he was equally unscrupulous?

10

Tax Incidence in a Market Economy

When a government levies a tax, the law assigns responsibility for payment to either a person or an institution. To understand the tax, however, one must look beyond the initial taxpayer. First, institutions cannot bear tax burdens; institutions are extensions of individuals, and only individuals can bear tax burdens. For example, the corporate income tax, while levied on an enterprise, ultimately may (1) reduce the profits earned by its owners, (2) raise the price of its products, and thus reduce the "real" income of consumers, or (3) lead to lower wages for workers.

Second, even if a tax is levied on an individual, its burden is not always borne by those who are initially responsible for paying. Directly or indirectly, its burden may be shifted to others. The *incidence* of a tax is the ultimate distribution of its burden.

In analyzing a potential change in tax policy, the first step must be to identify those who will benefit from the change and those who will be hurt and the magnitude of potential gains and losses. Changes in tax policy nearly always affect behavior, and when behavior changes, burdens are shifted around. This chapter uses the tools and concepts developed thus far to analyze the incidence of two taxes: the payroll tax and the corporate profits tax.

The Incidence of a Tax on Wages (Payroll) in a Competitive Labor Market

First, recall the factors that influence supply and demand in competitive labor markets. The amount of labor supplied at any wage rate depends on decisions made by households. Potential workers make choices in which they trade off the value that they place on leisure against the

value of the things that wages from working will buy. Judgments are made in accordance with individual preferences.

The shape of the supply of labor curve for any individual (and for the market) depends on the relative strength of *income and substitution effects*. When wages rise, two potential impacts on labor supply work in opposite directions. First, higher wages mean that for the same amount of effort workers will earn higher incomes. If we think of leisure as a normal "good," workers may "buy" more of it by *working less*. Recall that the wage can be thought of as the price of leisure. This is the *income effect* of the wage increase.

Second, however, a higher wage means that leisure is more expensive relative to other goods, and that workers may substitute other goods for it. That would imply *working more*. This is called the substitution effect of the wage increase. An upward sloping labor supply curve suggests that the substitution effect dominates the income effect. If the opposite were true, labor supply would "bend back."

In any event, the labor supply curve represents the behavior of workers with respect to the wage rate. Their behavior is a function of the wage that they actually take home *after tax* per hour of work.

The demand for labor depends on its productivity. Specifically, we determined that a profit-maximizing firm would hire labor up to the point that the wage rate was equal to labor's marginal revenue product (*MRP*). The marginal revenue product of labor is the revenue that the last unit of labor hired generates for the firm. The marginal product depends on technology (how much physical output an additional worker will produce) and output price (how much that output is worth).

Labor demand is, of course, a function of the full amount that firms must pay per unit of labor. That amount may include a tax if it is levied directly on payroll.

Figure 10.1 shows equilibrium in a hypothetical labor market with no tax on payroll. The equilibrium wage, W_0, is simultaneously the amount that firms pay and the amount that workers receive from unit of labor. Thus, the market will clear where supply and demand intersect. Although the supply curve is drawn with a positive slope, nothing should be inferred about the actual shape of the labor supply curve in any real society. We briefly discuss empirical estimates of labor supply elasticities below.

Now suppose that the government levies a tax of $\$T$ per unit of labor to be paid by employers. Figure 10.2 shows a new curve that is simply the old supply curve shifted up by an amount T. The new curve, labeled S', shows labor supply as a function of what firms pay. Regardless of whether the tax is paid by firms or workers, there is now a difference between what firms pay and workers receive. If workers receive W_0, firms must by paying $W_0 + T$.

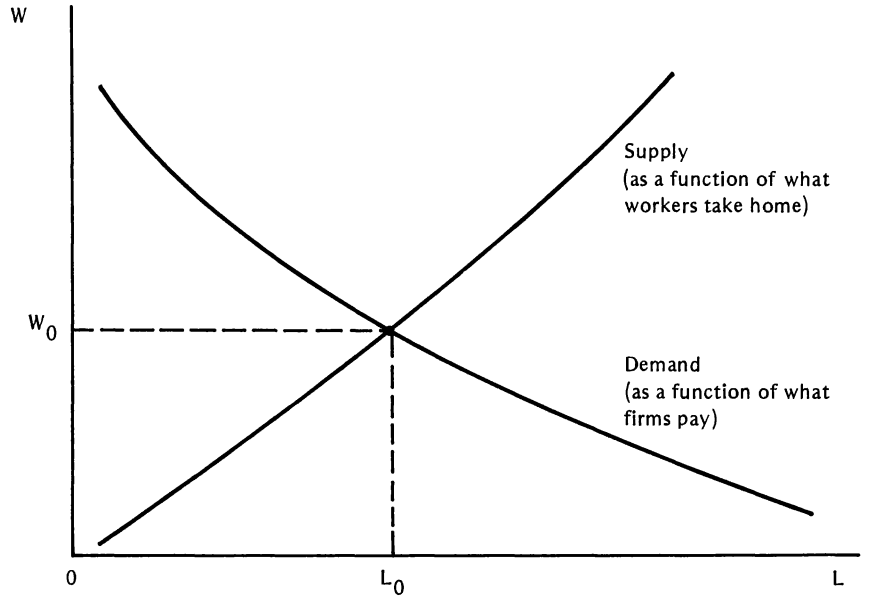


Figure 10.1. *Equilibrium in a competitive labor market—no taxes.*

If the initial wage was W_0, firms will pay $W_0 + T$ when the tax is first levied. But the higher cost per unit of labor will reduce labor demand from L_0 to Ld, and people will be laid off. At a gross wage of $W_0 + T$ there is an excess supply of labor.

With a surplus ($L_0 - Ld$) wages will fall. The issue, of course, is how far. Figure 10.2 reveals that a new equilibrium is achieved at W_1 with firms paying $W_1 + T$. When workers take home W_1, they will supply L_1; if firms pay $W_1 + T$, they will demand L_1, and the market clears.

The result is that the burden is shared by employers and employees. Initially, firms paid W_0; they now pay $W_1 + T$. Initially workers received W_0; they now receive W_1. Total tax collections are $T \times L_1$; geometrically it is equal to the whole shaded area in Figure 10.2. Workers' share is the lower portion $(W_0 - W_1)L_1$; the firms' share is the upper portion $(W_1 + T - W_0)L_1$.

Figures 10.3a and 10.3b show that the ultimate burden depends, at least in part, on the elasticity of labor supply. If labor supply is very elastic or responsive, most of the burden is borne by firms. If labor supply is inelastic or nonresponsive, most of the burden is borne by workers. The empirical work on labor supply in the U.S. economy suggests that for most of the work force, elasticity is quite low. The

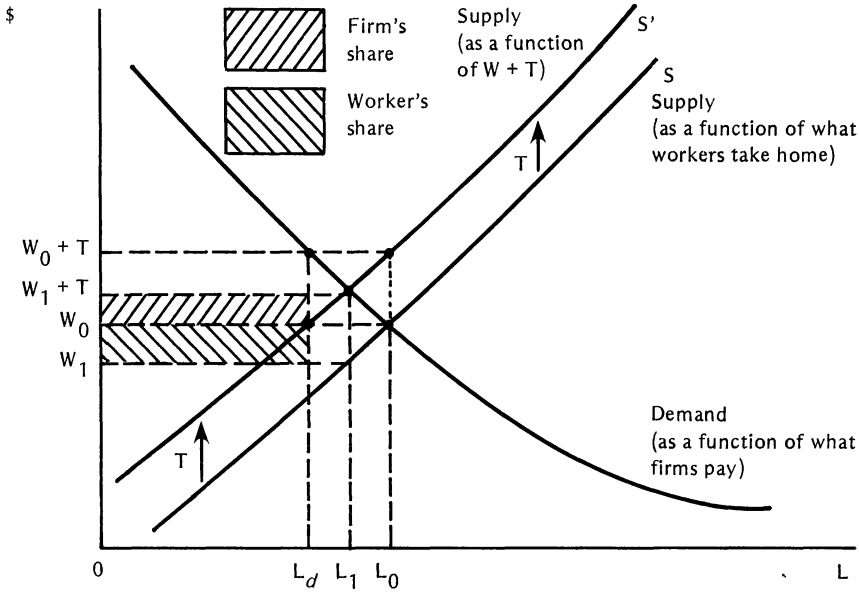


Figure 10.2. *Incidence of a per unit payroll tax in a competitive labor market.*

conclusion is that the bulk of the payroll tax is probably borne by workers.

The analysis produces the same result if the tax were initially levied on workers rather than firms. Using the same diagram but slightly different story, prove to yourself that this is true.

The Incidence of Tax on Capital (Profits) in a Two-Sector Model of the Economy (the Harberger Model)

We now shift to a tax on capital and complicate the story by making it a tax on one sector of the economy. Assume that there are only two sectors producing products X and Y. Assume further that there are only two factors of production: capital, K, and labor, L. Both K and Y are produced with capital and labor that are substitutable. In this chapter we abstract from the complexity of the capital market and simply think of firms as if they pay households directly for the use of capital as well as for labor (which is the end result of all that complexity, anyway).

Households bear tax burdens in basically two ways. Taxes may reduce their net income, or taxes may increase the prices of the things they buy. One ought to be indifferent about whether a tax doubles the price of all commodities but leaves income unchanged or whether a

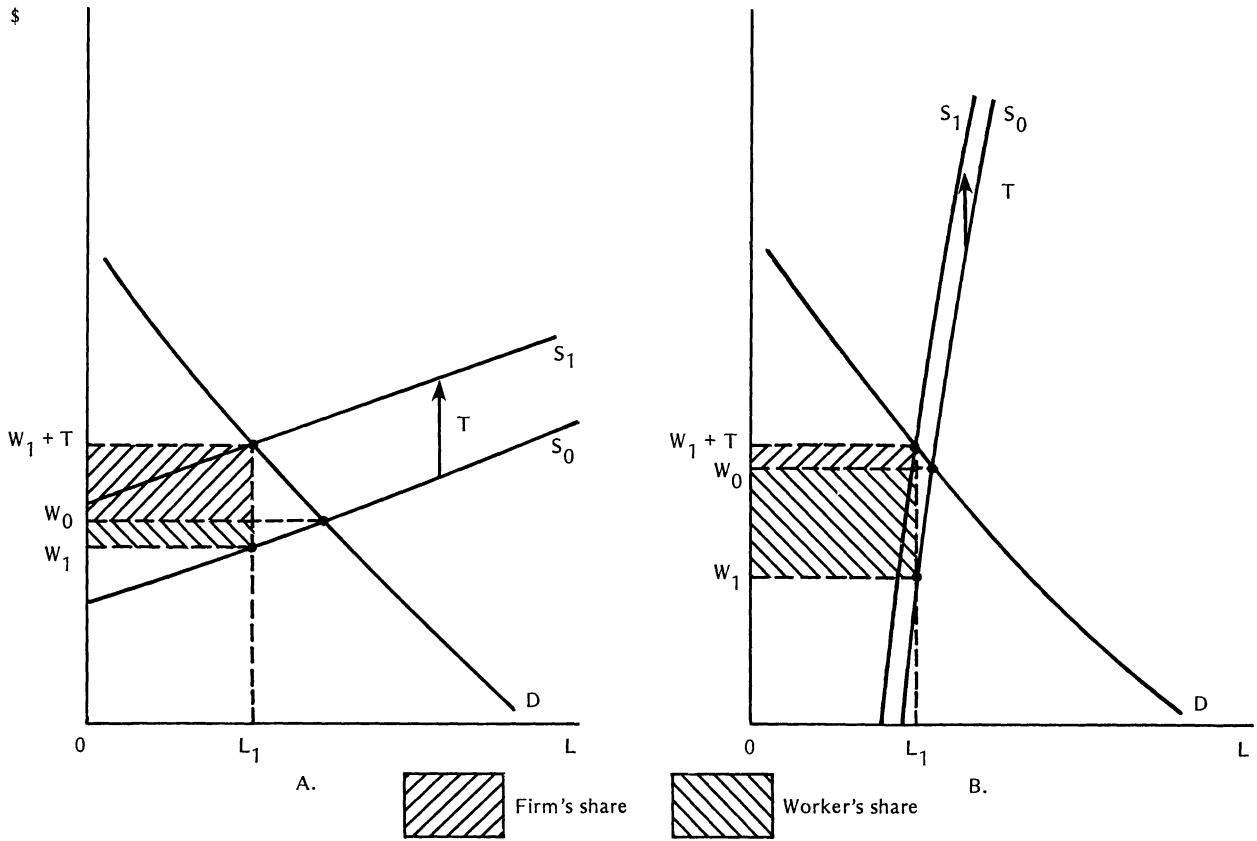


Figure 10.3. *Payroll tax with elastic (A) and inelastic (B) labor supply.*

tax cuts income in half but leaves prices unchanged. In the model, there are four prices: *Px, Py, Pk,* and P_l. P_l and *Pk* determine the incomes of laborers and capitalists respectively. P_X and P_Y are the prices of the traded goods. Burden can be reflected in changes in any one of the four.

One final assumption is that capital and labor are fully employed in the economy. That is, $L_X + L_Y = L^*$ and $K_X + K_Y = K^*$, where L^* and K^* are the total quantities of capital and labor in the economy. If demand for labor were to fall, wages would fall until labor was again fully employed. It should be noted that this is a *long-run* model, since capital flows from one sector to the other in response to changes.

The imposition of a tax, t_k, on capital in sector X would have several effects. First, the higher capital price ($Pk\,(1 + t_k)$) would provide firms in X with an incentive to substitute labor for capital. This would reduce the demand for capital in the market, and the price of capital, *Pk*, would begin to fall. Abstracting from any change in output for a moment, the demand for labor would rise, pushing up the wage rate, *Pl*. This is referred to as the *substitution effect* of the tax, and it will always work to the disadvantage of the taxed factor.

The size of the substitution effect depends on how easy it is for the two sectors to substitute labor for capital. Ease of substitution in production is measured by something called the *elasticity of substitution*. The elasticity of substitution in a sector is defined as

$$s = \frac{\text{Percentage change in K/L}}{\text{Percentage change in } P_l/P_k}$$

That is the percentage change in the capital labor ratio divided by the percentage change in the factor price ratio. An elasticity of 0 would indicate *no* substitution possibilities in response to factor price changes. A high elasticity of substitution would imply that it is easy to substitute the factors for each other.

A *high* elasticity of substitution in the taxed sector would indicate ease of substitutability and suggest that the impact of the tax on P_l and P_k is likely to be large. A low or 0 elasticity of substitution in the taxed sector would imply a *small* substitution effect.

The opposite relationship holds for the elasticity of substitution in the nontaxed sector. As firms in X reduce their demand for capital, *Pk* falls and firms in Y will hire more. Similarly, when firms in X increase their demand for labor, P_l rises and firms in Y will hire less. If it is *easy* for sector Y to substitute, only small changes in P_l and P_k will be needed to induce firms in Y to absorb the surplus K and release enough L to satisfy firms in X.

Next, there are market adjustments that result simply from the higher cost now associated with production in sector X. When costs in X increase, firms will cut back on output and supply less. In the long run (and this is a long-run story) firms will go out of business and leave X. Capital will flow to sector Y, which will expand. Initially, as the taxed sector faces higher costs and shrinks, the demand for capital and labor falls. The unemployed capital and labor cause P_l and P_k to fall. Firms in Y, seeing the lower factor prices, find it advantageous to expand and begin to absorb the surplus. Again, remember that this is the long-run model and that capital and labor markets must clear.

As factors are released in X and reabsorbed in Y, relative factor prices will change. If X is more capital intensive than Y, the relative price of capital will fall. If Y is the more capital-intensive sector, the relative price of labor will fall. This is called the *output effect* of the tax change; in general it works against the factor used most intensively in the taxed sector.

Finally, the reduction in supply in the taxed sector (X) and the increase in supply in the nontaxed sector (Y) will cause the price of X to rise relative to the price of Y. Thus, those who consume larger quantities of X will be hurt. The impact of the tax on output prices is called the *excise effect* of the tax change. Its size depends on elasticities of demand in both sectors.

Demand elasticities also determine the size of output effects. If demand for the taxed product is very inelastic, a small shift in supply will cause large excise effects, but quantity demanded and thus final output will not change very much. In essence, with inelastic demand, the tax is largely passed on to consumers. Substitution, output, and excise effects are summarized in Table 10.1.

In a now-classic article, Arnold Harberger used this model to analyze the corporation income tax in the United States (see Arnold Harberger, "The Incidence of the Corporate Income Tax," *Journal of Political Economy*, 1962). He used some rough estimates of demand elasticities, substitution elasticities, and capital labor ratios and concluded that the tax was fully borne by *capital income earners* in the economy as a whole (not just those who own corporations). The excise effects offset one another, since there is no reason to expect any one group to favor the consumption of corporate products.

The Property Tax

In most places the property tax is levied on the combined value of land and capital. In some countries (like Taiwan) it is levied on the value

Table 10.1
Summary of Two-Sector Model Adjustments to a Tax Imposed in Sector X

Effect	*Adjustments*	*Size Depends Upon:*
Substitution Effect	Works Against Taxed Factor Tax on K → P_K Falls Tax on L → P_L Falls	Elasticity of substitution in taxed sector (X) and non-taxed sector (Y) σ_X high → large impact σ_X low → small impact σ_Y high → small impact σ_Y low → large impact
Output Effect	Works against factor used intensively in the taxed sector If $\frac{K_X}{L_X} > \frac{K_Y}{L_Y}$ → P_K Falls Relatively If $\frac{K_X}{L_X} < \frac{K_Y}{L_Y}$ → P_L Falls Relatively	Elasticity of Demand for taxed good (X) η_x high → large effect η_x low → small effect
Excise Effects	Price of the taxed good (P_x) will rise Price of non-taxed (P_Y) good will fall.	Elasticities of demand for both goods

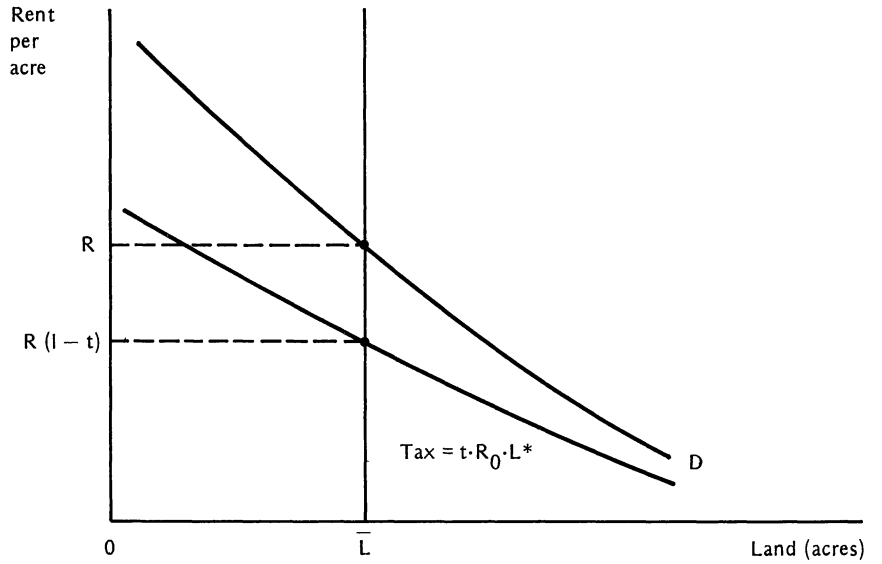


Figure 10.4. *Property tax on land.*

of land alone. We will first discuss the incidence of a land value tax and then turn to a more general property tax. Finally we will discuss the problem of differences in rates across local jurisdictions.

In the last chapter it was argued that since land is in fixed supply, its value is essentially demand determined. That is, each potential user implicitly calculates the return that the land will likely generate over time and bids accordingly. Land will, in the absence of restrictions, end up in its best alternative use earning a pure rent.

A tax placed on land value alone at an equal rate across uses will simply reduce the return to each proportionately and thus should not change the "rank order " of the bids. The winning bid should be the same as it was before but lower, and the allocation of land among alternative uses should not change. Since no economic decisions are changed, such a tax creates no *excess burden* (see Chapter 11) and cannot be shifted.

A property tax on land alone thus will be borne entirely by land owners who will see the net value of their land fall by the amount of the tax. This can be seen in Figure 10.4. There are no opportunities for shifting.

If the property tax were a national tax on the total value of capital and land at a constant rate across uses, the result for land would be the same. The incidence of the portion levied on capital (improvements) depends on what we assume about the elasticity of capital supply.

The Harberger two-sector model above assumed that capital was in fixed supply in the economy as a whole. Empirical estimates generally suggest that the savings rate (supply of capital) is not sensitive to changes in the interest rate. If we assume a fixed stock the portion of the property tax levied on capital will reduce the net (after-tax) rate of return across all uses by the amount of the tax. The result is that the burden will be borne by capital income earners.

It is important to recall from the last chapter that the value of a capital asset or of land is the present value of the stream of earnings to which the owner is entitled by virtue of his ownership. An asset (or piece of land) that will yield $5,000 per year for a long time will be worth $50,000 if the interest rate is 10 percent (paying less would imply a higher yield). Thus, a tax on the value of the asset can easily be translated into a tax on the income derived from it. A 2 percent property tax on a $50,000 asset payable annually is a flow of $1,000 per year. If the asset were yielding $5,000 per year before tax, such a property tax is equivalent to a 20 percent income tax on the yield.

If the supply of capital were elastic, the rate of capital accumulation would slow. Since the accumulation of capital raises the productivity of labor and thus pushes up wages in the long run, a tax that reduced capital supply would reduce wages in the long run. As a result, a portion of the burden would be shifted to labor. The evidence on capital supply suggests that this is an unlikely outcome.

In the United States, the property tax is locally administered in most states, which results in different rates across jurisdictions. This makes the problem of analyzing the incidence of the tax significantly more difficult.

Consider two jurisdictions identical in every way except for the property tax rate. Let jurisdiction *A* have the higher rate and jurisdiction *B* the lower rate. First, land cannot move, and to the extent that the tax is levied on land, land values will fall. Thus, land owners in *A* will bear a larger burden than land owners in *B*.

Capital, however, is mobile. Investment will favor jurisdiction *B*. As it does, land values in *B* will rise, and land values in *A* will fall further. Land owners in *A* thus bear part of the burden of the tax on capital as well.

If labor is fully mobile, employment will simply shift to *B*, and wages will remain equal in the two jurisdictions. If however, workers are "attached" to their initial jurisdictions, unemployment will push down wages in the high-tax jurisdiction (*A*), and wages in the low-tax jurisdiction (*B*) will rise. Thus, with immobile labor, workers in the high-tax jurisdiction will bear part of the burden.

Finally, the prices of goods produced in the two jurisdictions may adjust. To the extent that there are locally produced goods being sold

in local markets, prices in A would rise, and prices in B would fall.

Thus, when differentials in property tax rates exist across jurisdictions, the problem of burden is more complicated. While land and capital still bear that portion of the rate common to all jurisdictions, relatively high-tax jurisdictions could see lower land values, lower wages, and higher prices.

Note that capital in an owner-occupied house is no different conceptually than capital in manufacturing firm. Homeowners earn a form of capital income called *imputed rent*. If I live in a house that I own, I do not have to pay rent; the "yield" on my investment comes in the form of "housing services" that I consume over time. Thus, when we discussed the tax as a tax on land and capital, we made no distinction between capital and land in residential or commercial/industrial use.

It is possible for the property tax to be further shifted if markets are regulated. Consider, for example, the tax on rental property in a jurisdiction that has imposed rent controls. If the rent control board sets rents by simply calculating costs (including taxes) and marking them up, tenants will clearly bear the burden of the tax by paying higher rents.

Review Questions and Exercises

1. "New York State provides firms with an *investment tax credit* that effectively reduces the price of capital facing New York companies. If the demand for products produced in New York were relatively inelastic and firms in New York had very high elasticities of substitution, such tax credits would be likely to increase employment in the state." Discuss whether the statement is true or false and why.
2. Assume that there are two sectors in the economy: housing (H) and other goods (X). Assume that housing services are produced with capital alone and no labor. Describe the adjustments that would take place if the government were to place a tax on profits (or capital use) in the housing sector. Who would end up bearing the burden of such a tax? Answer the question assuming that a *subsidy* is paid to capital in housing production rather than a tax. In both answers use a two-sector model to analyze adjustments that follow.
3. A competitive industry (X) is in long-run competitive equilibrium. Assume that (a) it produces with two variable factors, K and L; (b) the elasticity of substitution between K and L is zero; (c) the capital labor ratio (K/L) is 1; (d) the elasticity of demand for X is -1; (e) it is a constant cost industry; (f) the industry is small

relative to capital and labor markets; (g) the price of capital is equal to the price of labor. Suppose that the government enacts an investment tax credit that has the effect of lowering the price of capital by 50 percent. Describe the response of the firm and the industry in the long run. Can you quantify the change in industry output that is likely to result in the long run?

11

The Efficiency of Competition: Sources of Market Failure and Optimal Tax Policy

The first ten chapters of this book have built and applied a model of a simple competitive economy. The logic is based on many implicit and explicit assumptions. It can be shown formally that if all the assumptions hold, the result would be an "efficient" overall allocation of resources. Recall that efficiency means that there are no changes that can be made in the final allocation of resources, mix of output, or distribution of output that would improve the well being of any without reducing the well being of others. Operationally, it means that no changes will result in value gains that exceed simultaneous value losses.

Unfortunately, however, all the assumptions do not hold, and unregulated markets do *not* produce an efficient overall allocation. There are many naturally occurring sources of market failure that can be corrected only through public sector or social involvement.

This chapter first presents an intuitive discussion of the formal proof of the efficiency of competition. It then introduces the notion of market failure and finally reviews the concept of efficiency as it relates to tax policy.

The Efficiency of Perfect Competition

Perhaps the most important part of the market efficiency logic is that a market system "produces what people want." Certainly, if personal preferences favor a set of goods that the system does not produce, the result would be inefficient. This notion of *consumer sovereignty* dates to

Adam Smith's description of the "invisible hand of competition"; it is today an important part of the formal logic of what is called welfare economics.

The condition that ensures that the "right" things are produced is

$$P = MC$$

That is, in the long run and short run a firm will produce where the price of its output is equal to its marginal cost of production.

The argument hinges on the proposition that in a free market, the price of a product is a good measure of the benefit that producing an additional unit of the product would yield to society. When a market price is posted and the market clears, anyone who subjectively values the good at or above its price simply buys it. Many of those people obtain a *consumer surplus*. That is, the product yields benefits to them in excess of its price; they would buy even at a higher price. Those who place a subjective value on the good that is less than its current price do not buy it.

If "one more" unit of a good were produced, price would presumably fall an infinitesimal amount and would induce a person to buy who before was indifferent to either buying or not buying. That person, when weighing the good against others, reveals that he values a unit of it precisely at *P*. Since we assume the good is private and that only one member of society derives benefits from any one unit, it is argued that the price is a good approximation of that good's *marginal social benefit*.

When firms equate price with marginal cost, they are then weighing the social benefit against additional cost. It remains to be argued that the marginal cost that a firm faces in its production decision is the full cost to society of the unit of the product. That would clearly not be the case if there were third-party costs (pollution, for example) for which the firm bore no responsibility. We assume for now that there are no such external costs.

To establish that marginal costs correctly measure the societal cost of added production, we turn to input markets. Resources are required to produce an added unit of output. Those resources will, through markets, come from one of two sources: previously unemployed resources or resources previously used to produce something else.

First, using labor as the example of resources that were previously employed producing something else, recall that workers are paid a wage just equal to marginal revenue product, which is the value of its product. If a unit of labor is attracted from one firm to another it is because its product is valued more highly there.

At equilibrium, with many firms each buying labor up to the point where $w = MRP_l$, the marginal value of a unit of labor's product in X is virtually the same as the marginal value of a unit of that labor's product in Y, and so forth. Thus, if all resources are valued in competitive markets, the marginal cost of a unit of any output is approximately equal to the value of the goods that the resources would otherwise have produced.

If, on the other hand, the resources had been previously unemployed, the logic still holds. Again, consider labor as an example. If the market clearing wage is equated by all firms to the value of labor's marginal product and an individual *chooses* not to be in the labor force, that individual has revealed that leisure is worth more to him than the value society places on his potential product. It is therefore efficient for him not to work. This, by the way, does *not* hold when markets do not clear, such as when there is involuntary unemployment.

If a previously unemployed worker "tips" over the line and decides to join the labor force, he will be giving up leisure that he values less than the wage. Those who value leisure more highly will not take a job; those who place a lower value on it are already working.

Marginal cost of production represents the cost of resources needed to produce one additional unit of a product. If resources are valued in competitive markets, their prices approximate their opportunity costs to society. If they come from producing something else, society loses products of approximately equal value; if they were previously unemployed, society gives up leisure of approximately equal value.

The condition that price equals marginal cost or $P = MC$ then implies that each firm (industry) would produce as long as the marginal benefit of its product to society (reflected in P) exceeds the value that society places on what must be given up in order to produce it (reflected in MC). If, however, price and marginal cost diverge for any reason, or if production generates external costs or benefits that the firm has no incentive to consider, the result will be inefficient.

Formal proof of the efficiency of perfect competition is, of course, much more technical. It involves a level of sophistication and analysis far beyond what can be presented here.

The Economics of Market Failure

When all the assumptions of competitive markets hold, the result is an efficient allocation of resources. The market will produce what people want and will produce it at least cost. No changes can, on balance, improve efficiency. Real market conditions, however, are not as nice. Many assumptions do not hold and others hold only in certain circum-

stances. There are four major sources of market imperfection that will be introduced and briefly discussed here: (1) imperfect market structure (noncompetitive behavior), (2) the existence of "public goods," (3) the presence of "external" costs or benefits, and (4) imperfect information.

Imperfect Markets

The operational assumption through Chapter 11 has been that in each market there is a large number of buyers and sellers. When each buyer and seller is small relative to the market, none can directly influence price; all economic decision makers take input and output prices as given.

While a Kansas wheat farmer is probably a "price taker," the Toyota Company is certainly not. A very large percentage of all firms have some control over price. In the case of a single-firm industry (monopoly), the only constraint on price setting is market demand. When other firms compete for market share, strategic pricing behavior may develop.

Many formal models have been developed that attempt to capture the likely behavior of price setting noncompetitive firms. A model of monopolistic behavior will be developed in more detail in Chapter 12. The result of nearly all these models is that price and marginal cost diverge. A profit-maximizing monopolist, for example, will raise price above the competitive level and cut output correspondingly. The result is substantial *underproduction* from society's viewpoint. Regulation or restructuring through antitrust enforcement may lead to a more efficient outcome.

Public Goods

A second major source of market failure is that private producers will not find it in their interest to produce everything that members of society want. More specifically, there is a whole class of goods called *public goods* or *social goods* that will be underproduced or not produced at all by an unregulated market economy.

Public goods are goods that bestow collective benefits on society. That is, they are in a sense collectively consumed. The classic example is national defense, but there are hundreds of others. Some public goods, such as national defense, provide benefits to the whole nation. Others, such as clean air, may be limited to smaller areas; the air may be clean in the suburbs but polluted in the central city.

The difficulty with public goods is that consumption cannot generally be prevented for nonpayment. Once the good is produced, no one can be selectively excluded from enjoying its benefits. Producers of pri-

vate goods can make a profit because they do not have to yield the product until you pay.

A voluntary payment system for public goods presents consumers with a dilemma. Even if I want the good, there is no contingency associated with payment; I get the good whether I pay or not, and my contribution is so small that it will not affect the level of output.

Thus, private provision fails, and members of society must get together to ensure that desired public goods are produced. This generally is accomplished through government provision and tax finance. Notice that the purpose of government provision is to correct for a naturally occurring failure of the market to produce what members of society desire.

Externalities

A third major market imperfection is the existence of external costs and benefits. By *external* we mean that there is a cost or benefit imposed or bestowed on someone or some group external to the transaction, a third party. In an urbanized society, external effects are pervasive; they are more the rule than the exception. The classic example is pollution, but there are others, such as airport noise, traffic accidents, a loud radio, congestion, and painting one's house an ugly color. Not all are negative; housing investment may yield benefits for neighbors.

Externalities are a problem if private decision makers do not account for them. The efficiency logic presented at the beginning of this chapter was based on firms' weighing social benefits off against social cost. If a firm in a competitive environment produced a good, it was because the value of the good to society exceeded the social cost of producing it. If social costs or benefits are overlooked, decisions may simply be wrong or inefficient.

The market itself provides no automatic mechanism that provides decision makers with an incentive to consider external effects. Society, however, has established over the years a number of different institutions for dealing with them. Tort law is a body of social rules that deals with third-party effects. Under certain circumstances, those who impose costs are held strictly liable for costs that they impose; in other circumstances liability is assessed only if the cost imposed resulted from negligent behavior.

Economists have for years suggested that a carefully designed set of taxes and subsidies could help "internalize" external effects. For example, if a polluter were taxed in proportion to the damage actually caused by her pollution, she would consider those costs in her decision to pollute.

Calculating damages in pecuniary terms is often a difficult task, but it is an unavoidable one. Judges in liability cases are forced to make such judgments all the time. Public policies to deal with problems like "acid rain" will hurt one sector at the expense of another, and unless absolute rights are involved, gains and losses must be weighed.

Imperfect Information

A fourth major source of market imperfection is imperfect information on the part of buyers or sellers. The conclusion that markets work efficiently rests heavily on the assumption that consumers and producers have full knowledge of the characteristics of products, available prices, and so forth. The absence of full information can, of course, lead to transactions that are ultimately disadvantageous.

Some products are so complex that consumers find it difficult to judge the potential benefits and costs of purchase. Certainly, demanders in the markets for medical care do not fully understand what they buy.

Some forms of misinformation can be corrected with simple rules such as "truth in advertising." Others, like the problem of medical care, are much more complex. Certainly there is a role for social or governmental action to ensure that buyers and sellers are as fully informed as possible about the qualities and prices of the things they exchange.

Tax Policy and Efficiency

When judging the desirability of a tax policy, the criterion of economic efficiency is an important one that should not be overlooked. Certainly there are many criteria that in any given circumstance will be used to judge the desirability of a change in tax policy: the impact of a change on the distribution of income and wealth, the perceived fairness of the outcome, the potential impact of the change on macrostability (unemployment then inflation), political considerations, and so forth. All must be weighed along with efficiency.

Taxes and tax policy may be used to correct an existing inefficiency, or taxes themselves may be the source of inefficiency. Examples of the use of tax policy to improve the performance of the market include the taxation of activities producing external costs. Taxation for purposes of public goods production can certainly be justified on efficiency grounds.

On the other hand, taxes themselves are often the source of inefficiency. When a tax distorts or alters economic decisions, it can be shown that the burden imposed on society is greater than the amount of tax collected. Since all taxes affect economic decisions to some extent, the

role of tax policy must be to minimize rather than eliminate these excess burdens.

Review Questions and Exercises

1. An excise tax that raises the price of a product to consumers, places a wedge between the price paid by the consumer and the marginal cost of production in the firm. Such a tax is inefficient. Can you explain why?
2. A paper factory dumps polluting chemicals into a river. Citizens who live along the river (there are thousands) bring a class action suit claiming damages. Briefly discuss the fairness and efficiency of the following possible rulings of the judge:
   a. Simply affirm the polluter's right to dump in the river;
   b. Find for the plaintiff—the polluter must compensate residents for any damages imposed;
   c. Refer the matter to the Environmental Protection Agency, which imposes a tax on the firm equal to marginal damage costs; proceeds are *not* paid to damaged parties;
   d. Order the firm to simply stop all pollution.
3. The following appeared in the *Boston Globe* on Monday, April 18, 1977:

   > New York—A caravan of about 550 cars eked its way through Kennedy Airport to protest the Supersonic Concorde Airliner yesterday, and its organizers claimed victory despite failure to clog the places as predicted.
   >
   > Those involved in the protest were demanding that the Port Authority of New York and New Jersey make permanent its year old ban against use of Kennedy by the Concorde. It put off a decision last week for a third time pending evaluation of tests of Concorde operations at Washington's Dulles Airport and in Europe. . . . People objected to the Concorde because they considered it noisy and they felt it caused vibration and air pollution.

   You are appointed as a consultant to the New York Port Authority to determine whether such a ban would be "efficient." What information would you need to make such a determination? If you decide that an outright ban would be "inefficient" but that allowing the Concorde to land at Kennedy would lower property values and thus be inequitable, what kind of resolution might you propose to the Authority?

4. Public goods can be produced by the central government or by local governments. Can you think of principles that you might use to assign particular public goods to particular levels of government?

12

Prices, Output, and Efficiency in Monopoly Markets

In Chapters 1 through 10 we developed a model of a simple economy in which there were many buyers and sellers interacting in each market. The opposite, of course, is a market in which there is only one seller: a monopoly. This chapter will explore in detail the decisions facing a single-firm industry. As before, we will discuss three fundamental decisions: output supply, choice of technology, and input demand. However, now we add a fourth: choice of output price.

There is no real difference between the behavior of a competitive firm and the behavior of a monopolist on the "cost" side of the profit equation. Both are assumed to be price takers in input markets. Both will choose the technology that produces output at least cost.

The difference arises on the revenue side of the ledger. A competitive firm faces a fixed-market determined price, and we assume that it can sell all it wants at that price, constrained only by its current capacity in the short run. The *demand* curve facing a competitive firm is simply a horizontal line. Marginal revenue to a competitor is simply the price of the output, since all *marginal units* are sold at the same price.

A monopolist, however, faces market demand, and the amount he is able to sell depends on the price he chooses. Once price becomes a choice variable, the nature of marginal revenue changes. We assume that the monopolist cannot *price discriminate*; that is, she must choose one price/quantity combination and charge every consumer the same price. Increasing output, then, involves not only producing one more unit and selling it, but also reducing price across the board.

Consider the hypothetical demand curve in Table 12.1 and Figure 12.1. Suppose that we wanted to increase our target output from three units to four units. We would sell the fourth unit for $7, but since we

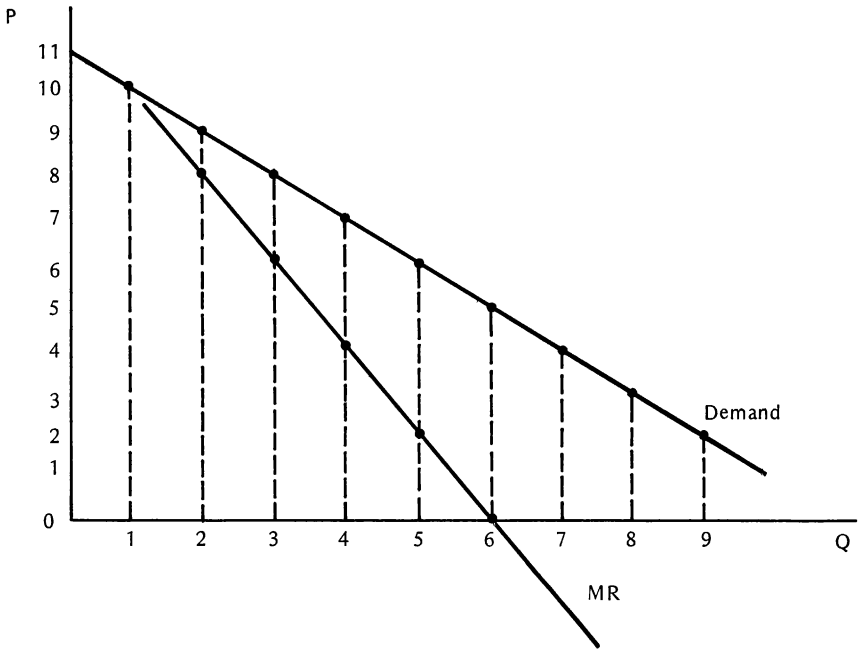


Figure 12.1. *Marginal revenue facing a monopolist.*

Table 12.1
Demand and Marginal Revenue Facing a Monopolist

Quantity	Price	Total Revenue	Marginal Revenue
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

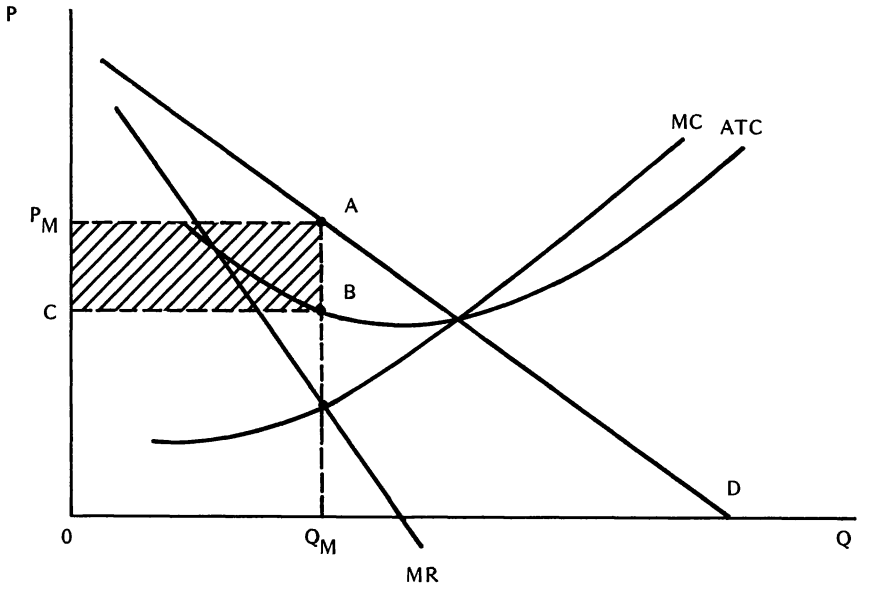


Figure 12.2. *Price and output choice for a profit-maximizing monopolist.*

cannot price discriminate, we must sell *all four* for \$7. Had we chosen to produce only three, we could have sold them for \$8 each. Thus, offsetting the \$7 of revenue gain is a \$3 loss, that is \$1 for each of the three units we could have sold at the higher price. Marginal revenue is thus only \$4, which is considerably below price (\$7).

Similarly, an increase in target output from four units to five units generates a marginal revenue of \$2. The fifth unit sells for \$6, but the first four would have sold for \$7 rather than \$6 if output had been limited to four. Thus, marginal revenue is \$6 minus \$4, or \$2.

For a *continuous straight-line demand curve*, the marginal revenue schedule lies below it and will bisect the X-axis between the origin and the X-intercept of the demand curve.

Figure 12.2 superimposes a demand curve and the marginal revenue curve derived from it on a standard set of cost curves. The basic decision process is now the same as it was for a competitor. A profit maximizer will raise his production target as long as the added revenue from the increase exceeds the added cost, thus enhancing profit. More specifically, as long as marginal revenue is greater than marginal cost, it is profitable to raise output; the difference between them can be thought of as *marginal profit*. The optimal output price combi-

nation for the monopolist in Figure 12.2 is *Qm* and *Pm*. At any output below *Qm*, the monopolist would raise output, since marginal revenue is above marginal cost; the firm would never produce above *Qm*, since marginal cost exceeds marginal revenue.

At *Qm*, price will be fixed at *Pm* yielding revenues of *Pm* x *Qm*, or the area *PmAQmO*. Total cost is the product of average total cost and *Qm* or the area *CBQmO*. Total profit is thus *ABCPm*.

In competition, the presence of economic profits would provide an incentive for firms to enter, and price would be competed down, eliminating profits. Often, however, monopolies are able to restrict competition and prevent the entry of competing firms. Examples of potential barriers to entry are an exclusive government license, a patent, a powerful brand name, or simple coercion.

The inefficiency and potential loss associated with monopoly can be seen by comparing the monopoly outcome with the competitive outcomes where there is no technological advantage to large size. Consider, for example, an industry with many firms operating with a production technology exhibiting constant returns to scale. Figure 12.3 is a diagram of the long-run competitive equilibrium of such an industry. Recall that the industry supply curve is the horizontal sum of all the individual short-run marginal cost curves.

Now suppose that the government nationalizes the industry and then turns over control to a single private monopolist. The monopolist owns a firm with many plants. Nothing technological has changed, just the locus of decision making power.

To analyze the monopolist's decisions we must derive the consolidated cost curves that the single firm now faces. In fact, the new firm's marginal cost and average cost curves are simply the horizontal sums of the cost curves of the individual plants (old firms). Figure 12.4 shows marginal and average cost curves for a three-plant firm. Since the firm owns all three plants, the average cost of producing 9,000 units is $12: 2,000 produced in plant 1 at an average cost of $12, 3,000 produced in plant 2 at an average cost of $12, and 4,000 produced in plant 3 at an average cost of $12. Similarly, the marginal cost of the 16,000th unit is $10 when 4,000 are being produced in plant 1, 5,000 in plant 2 and 7,000 in plant 3.

We can now superimpose the cost curves of the consolidated monopoly industry on the diagram of the competitive industry in Figure 12.3 (see Figure 12.5). The demand curve will pass through the minimum point on the consolidated firm's average total cost curve. (Convince yourself that this is true.) The old competitive supply curve was simply the sum of the individual firm marginal cost schedules, but recall that the horizontal sum of the marginal cost curves is also the marginal cost curve of the consolidated firm.

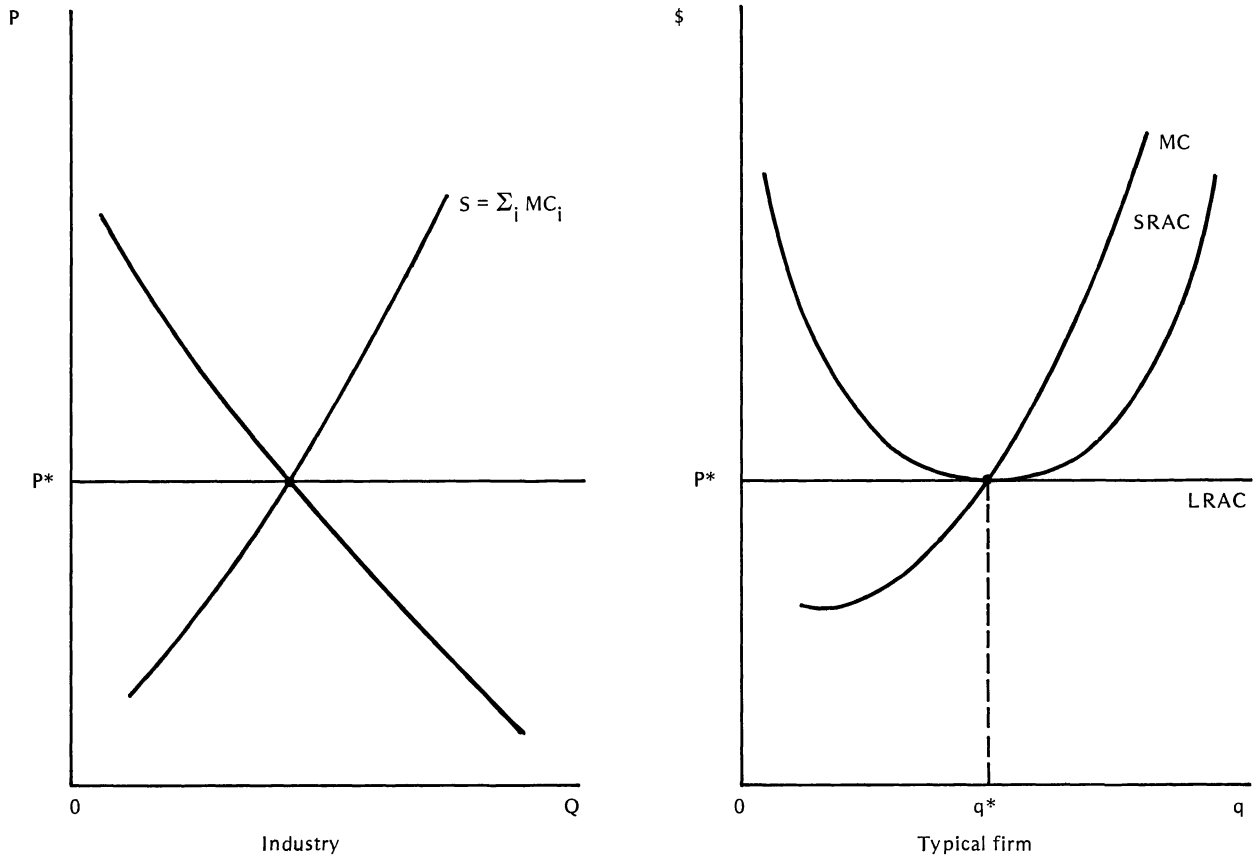


Figure 12.3. *A competitive industry in long-run equilibrium.*

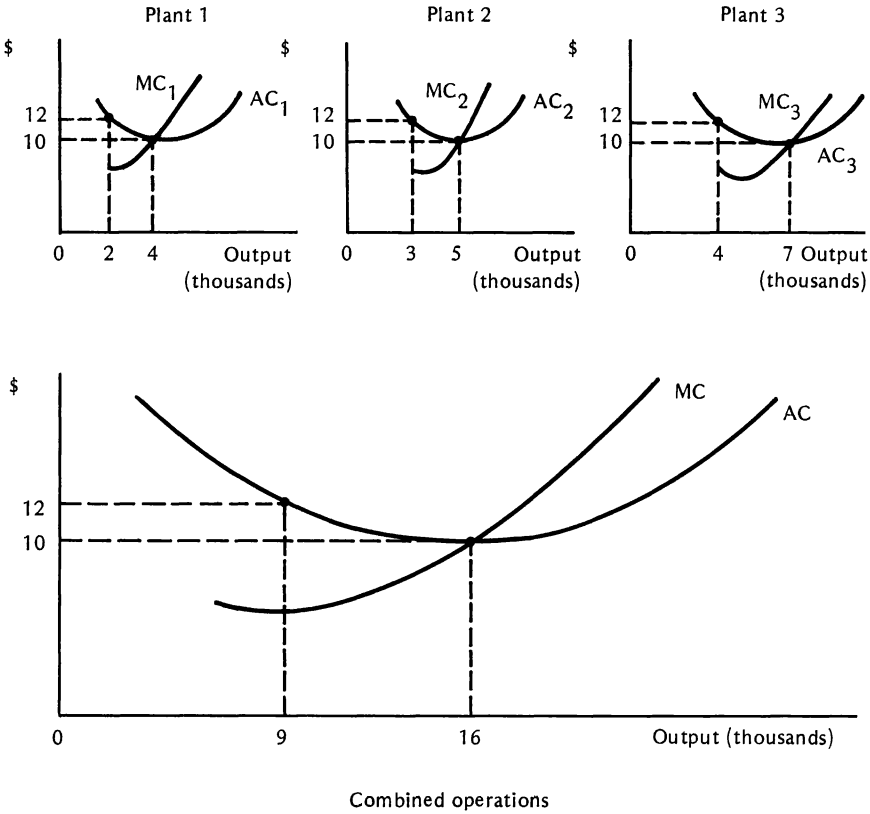


Figure 12.4. *Cost curves for a consolidated multiplant firm.*

The new monopolist will no longer produce where demand intersects the old supply schedule, which is at *Pc* and *Qc*. It will produce where marginal cost and marginal revenue intersect, at *Qm*, and it will charge *Pm*. The final result is that output is reduced, price is raised, and price at equilibrium is substantially above marginal cost.

The Welfare Cost of Monopoly

It was argued in Chapter 11 that the condition $P = MC$ is central to the proposition that markets produce "what people want" and that a competitive allocation of resources is efficient. We now note that a profit maximizing monopolist will produce where price is greater than marginal cost. The simple argument was that price is a good approximation of the social value of a unit of output and that marginal cost, in

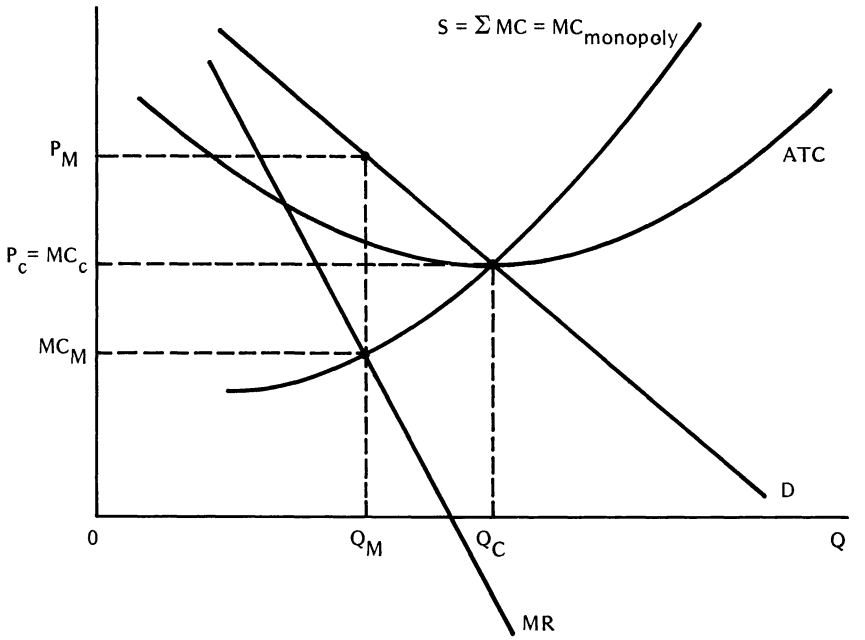


Figure 12.5. *Comparison of monopoly and competitive outcome for a firm with constant return to scale.*

the absence of externalities, is a good approximation of the social opportunity cost of the good. If price is greater than marginal cost, then additional production should yield a *social dividend.*

Figure 12.6 presents a slightly simplified version of the earlier monopoly diagram. Here for clarity we ignore the short-run cost curves and assume constant returns to scale in the long run. Under competitive conditions firms would produce *Qc* and price at *Pc* equal to *MC.* Under monopoly, the firm would produce at *Qm* and price at *Pm.*

Consider a shift from the competitive output to the monopoly output. Previously, consumers enjoyed a *consumer surplus* equal to the area of triangle *DBPc*, and firms earned *0* economic profits. The cost of resources was *Pc* x *Qc*, and it was fully borne by consumers. After the shift there are some gains and some losses. The monopolist, of course, now earns an economic profit of *PmACPc.* Consumers are the losers, since they must now pay a higher price for the same good.

Consumers who now pay *Pm* find their surplus reduced to *PmDA*, a loss of the area *PmABPc.* Thus, the loss to consumers exceeds the gain to the monopolist, and the result is inefficient. The net loss to society can be approximated by the area of triangle *ABC.*

There are several remedies available to deal with monopoly power. The U.S. antitrust laws favor "the competitive solution." Large firms

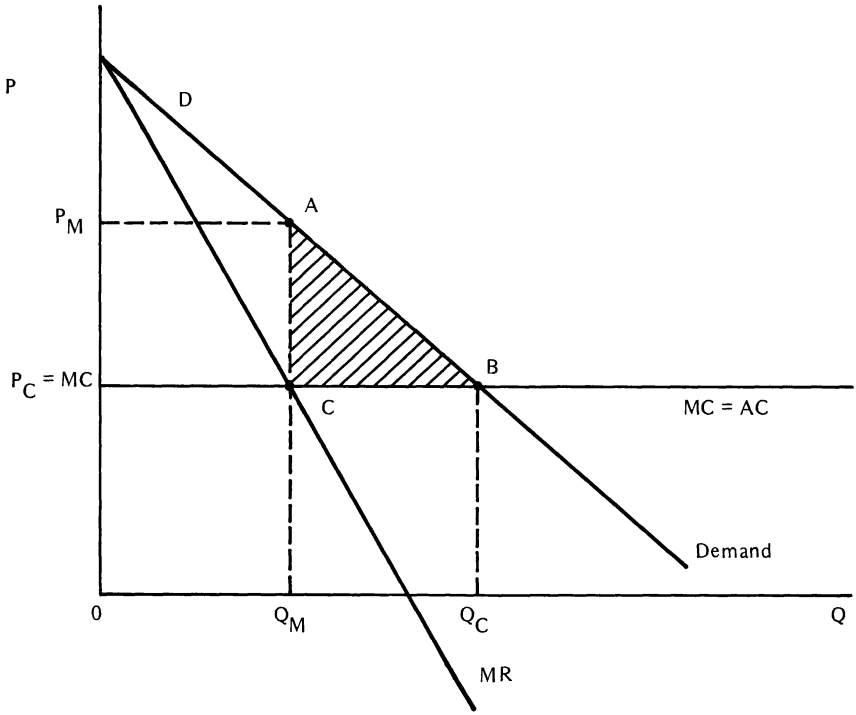


Figure 12.6. *Welfare loss from monopoly.*

have been broken into smaller competing firms on many occasions. The most recent example is the split up of the American Telephone and Telegraph Company. Earlier in this century, The American Tobacco Trust and Standard Oil Company were carved up into competing pieces.

Alternatively, monopolies might be regulated. A simple price ceiling placed on a monopoly at marginal cost would provide firms with an incentive to actually increase output. The U.S. courts and the U.S. government have been reluctant to use regulation as an antitrust remedy except when the firm is a "natural monopoly."

Natural Monopoly

The example above describes an industry in which the efficient scale of operation is small. There was no technological reason to have big firms instead of small firms. While examples are rare, it is possible that the technological economies of scale of an industry are so large that it makes sense to have just one large firm. An example often cited

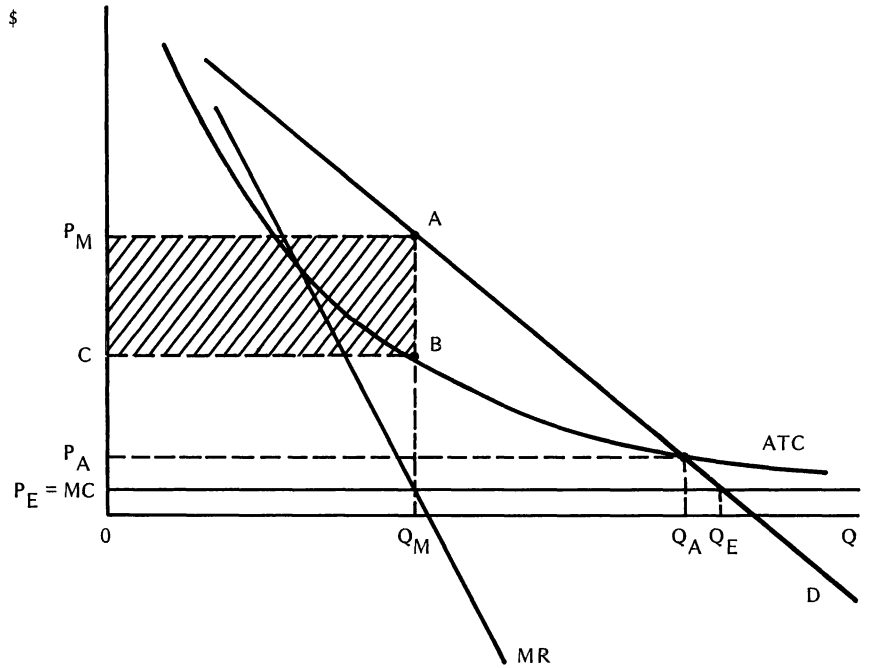


Figure 12.7. *Natural monopoly.*

is a public utility. A firm that has such a cost structure is called a *natural monopoly*.

A diagram of the cost structure of a natural monopoly is presented in Figure 12.7. We assume for simplicity that the firm has some very large initial fixed costs and very low constant marginal costs (*MC*). Notice that demand is fully exhausted while average cost is still declining. To operate such a business on very small scale would be silly, since costs of production rise rapidly as firm size is reduced.

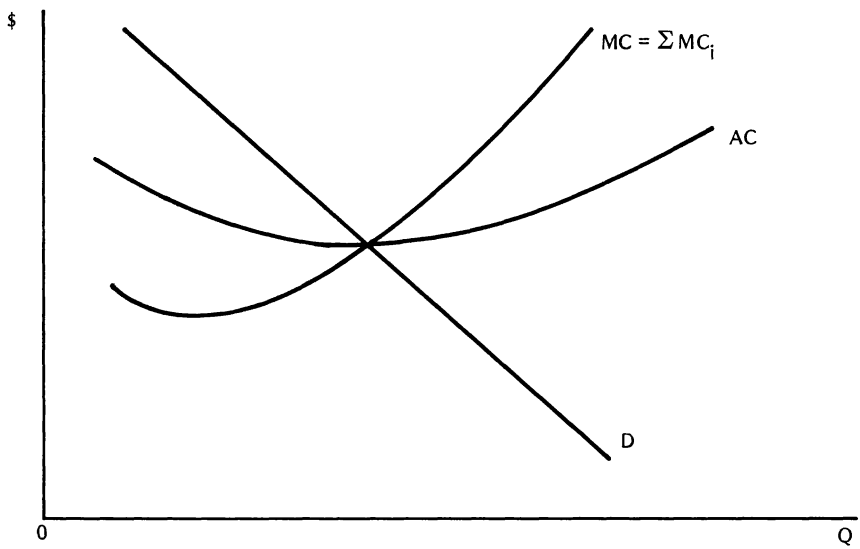
If an exclusive government license were issued to a single firm that was then allowed to choose price and quantity, it would price at *Pm*, produce at *Qm*, and earn economic profits of *PmABC*.

The efficient price would be $P_E = MC$. If a price ceiling were imposed at that price, however, the firm would suffer losses and go out of business; price would be below average cost.

There are two possible solutions. The best would be to charge P_E and to subsidize the business out of nondistorting tax revenues. Except for any distortions or inefficiencies associated with the tax, this is the efficient policy. Most public utility commissions allow firms to charge a price, P_A, that covers average costs including a "normal" rate of return to owners of the firm. While still inefficient, it represents something of a compromise.

Review Questions and Exercises

1. Assume that the potato chip industry in Massachusetts in 1975 was competitively structured and in long-run equilibrium earning a normal rate of return. In 1976 two smart lawyers quietly bought up *all* the potato chip firms and began operations as a monopoly called "Wonks." In order to operate efficiently, Wonks hired a management consulting firm that estimated long-run costs and demand. These results are presented in the diagram below:



   a. Indicate 1975 price and industry output.
   b. Assuming the monopolist is a profit maximizer, indicate total revenue, total cost, and total profit *after* the consolidation.
   c. Assuming that about 25 percent of potatoes grown in Maine are used for potato chips in Massachusetts, what would you expect to happen to the price of Maine potatoes in the short run after the consolidation? Suppose that after three years of successful operation, the *annual* accounting statement looked like the following:

Price per bag	$.50
Total sales (bags)	$4,000,000
Cost of potatoes	$1,000,000
Total additional cost (including labor, maintenance, depreciation, etc)	$500,000

General Foods in 1978 notices Wonks's success and begins researching a possible takeover. Its research indicates that

a. A "brand name" and several patents on production techniques serve as an effective barrier to entry;

b. Forecasts indicate constant demand and costs for many years to come.

d. How much would General Foods probably be willing to pay for 100 percent of Wonks stock if AAA Bonds were presently yielding 10 percent?

e. The following year, an old law school buddy files a complaint with an antitrust division of the Justice Department claiming that General Foods now controls 100 percent of the New England potato chip market. Justice concurs and prepares a civil suit. You are asked to write a *brief memo* (two to three paragraphs) to the new president citing

   a. The economic justification for action;

   b. A proposal to achieve an efficient market outcome;

   c. The probable effects on the "value" of General Foods subsidiary, Wonks.

13

Summary and Conclusion

It is essential for government officials, particularly planners and tax policy makers, to understand the likely effects of their actions on economic systems. If tax laws are changed, who will benefit and who will lose? If we switch from a sales tax to heavier reliance on the property tax or to a land tax, how will that alter the composition of output? How much of a burden will it impose on society?

In addition, policy officials in every country continuously struggle with the issue of the proper role of government. Competing views on the relative importance of government and the market in determining economic outcomes lie at the heart of international politics. What decisions can be left to the market? In what circumstances do markets seem to work well? When do they break down? What is the *proper* role of tax policy?

The Basic Logic of the Market System

This book has discussed both positive and normative economics. We began by building a simple model of a competitive market economy. The model of *perfect competition* serves two purposes. First, even though it involves a number of simplifying assumptions, it captures the essence of the *market allocation mechanism* very well. The argument is that one cannot hope to understand complicated markets without understanding simple ones first.

Second, it serves as a useful point of departure for a discussion of the role of government. Markets seem to do some things very well, but when the competitive assumptions are relaxed, major sources of naturally occurring market failure are exposed. Economic theory does

not suggest that a laissez-faire economic system will lead to an efficient or equitable allocation of resources. Rather, it suggests several very specific roles for public involvement in the allocative process.

Simple supply and demand diagrams are useful for showing the basic logic of the price system. Final goods and services as well as inputs are rationed to the highest bidder. That is, those who are willing and able to pay the most are the ones who get the goods. On the resource side that implies that inputs end up in their most profitable uses.

One of the basic assumptions behind the concept of economic efficiency is that economic systems should produce what people want. The logic is that people themselves are the best judges of what does and does not improve their well-being. *De gustibus non est disputandum*: You can't quarrel with tastes.

A simple competitive market system responds to preferences no matter how bizarre without central direction. Consider a system at long-run equilibrium. If consumers decide that they want more of some good, X, that will be expressed in the market; demand will shift. An increase in demand will create a shortage, and the price of X will rise. As it does firms producing X begin to earn higher profits. If they were at long-run equilibrium before, they will now be earning economic profits, profits over and above a normal return to invested capital.

Those higher returns provide existing firms with an incentive to expand and an incentive for new firms to enter. As capital flows in, the demands for inputs required to produce X shift. The prices of those inputs rise, and they will be competed away from alternative uses. If there are skills specific to the production of X, the return to acquiring those skills will rise providing those entering the labor market with an incentive to get the appropriate training.

The market system also provides firms with the incentive to produce goods at minimum cost to society. The operational assumption is that firms maximize profits; firms that do not will presumably be competed out of business. Since profit is the simple difference between revenues and costs, firms must minimize costs.

None of this requires central planning or control. In modern industrial societies, production processes are complex. The knowledge required to set up and operate a single firm may be enormous. In a market system, decentralized decision making limits the amount of information that any single decision maker must digest. No central agency must understand the production technologies of all industries.

The efficiency logic of the market system can be found in the competitive market equilibrium conditions:

$$P = MC$$

$$w = MRP_l$$

$$r = MRP_k$$

Profit-maximizing firms will produce output up to the point that the price of output is equal to the marginal cost of production. In a market system, where people are free to buy or not to buy each product offered for sale, price is a good reflection of each product's *marginal social value*. Those who value a product more highly than its current price buy it; those who personally value it less than its price do not. Thus if one additional unit of a good were produced, it would go to the person who just values it at its price; it would add to supply on the market driving price down enough (presumably a tiny amount) to induce one person to buy that unit.

Since resources are valued in competitive markets, marginal cost is a good approximation of a good's marginal social cost. The appropriate measure of social cost is the value that society gives up to produce this good. When we increase production of a good, resources are required, and they must come from an alternative use. As an example, consider labor.

If labor is needed to produce more X, its alternative use was either leisure or production of some other good. If labor markets are open, workers compare the value that they place on leisure with the prevailing wage. That wage in turn reflects the value of labor's product. At equilibrium in competitive markets we know that $w = MRP_l$. MRP_l is the market value of the product that a single unit of labor will produce in each market.

Thus, workers facing the market wage weigh the value that they place on leisure against the value that society places on the product that they would produce on the job. Firms in turn compare the value that society places on their product with a cost that reflects the real social value of the inputs used.

The result is an efficient allocation of resources. Firms produce the goods and services that people want, and they have an incentive to do so at least cost.

Market Failure and the Role of Government

All of this, however, rests on a set of assumptions that do not hold in real economies. As we begin to relax them, a number of serious problems expose themselves; they suggest a number of potentially important roles for social or government regulation of the market.

First, there are a set of things that members of society might want that the private market does not have an incentive to provide. These are called *public goods*. The problem with most public goods is that people cannot be "excluded" from consuming them for not paying.

The private market works to produce the appropriate *private* goods because people must "reveal" their preferences by buying them. If we don't pay, we don't get the hamburger. If benefits are collective, however, the contingency is gone. Once police protection and a system of courts reduces the crime rate, we enjoy a safer society whether we pay or not.

Unless people are completely altruistic, a private producer will be unable to collect voluntary contributions for providing public goods. A self-interested person will find no benefits linked to paying. Her contribution will be very small relative to the amount needed to produce the good, and she cannot be excluded for nonpayment.

The nature of public goods may also lead to dissatisfaction with the private sector. Since benefits are collective, a *single* level of production must be chosen that we all in a sense consume. In private markets we all face a set of prices, and we simply can choose for ourselves how much to consume. If society chooses a level of public goods production that is consistant with the preferences of "the median voter," half the people will think that it is too much and half the people will think that it is too little.

A second major source of market failure is the existence of externalities. Often private producers and consumers do not naturally weigh all the benefits and costs involved in their decisions. The most-often-cited example is environmental damage. Industries in Ohio and other midwestern states pump waste products into the air that are carried by the prevailing winds to the east, where they fall from the sky in the form of acid rain. Some estimates of the costs involved are enormous; acid rain may be killing forests and fish in eastern lakes as well as causing serious damage to buildings. If such costs do not figure in the decisions of producers, the result may be a serious misallocation of resources.

In urban societies where people live in close proximity to one another, external effects are everywhere; we make few decisions that do not affect others. For an efficient allocation of resources, society must ensure that these externalities are internalized: That is, decision makers must have an incentive to consider them.

A number of mechanisms have been used to this end. First, all societies have laws that define rights and impose penalties when they are violated. For example, liability rules impose costs on those who cause accidents by being negligent.

Second, tax and expenditure policy may play an important role. Activities generating external costs might be taxed, and if the tax is set to properly reflect the true external costs resulting from the activity, the result will be efficient. Similarly, activities that generate external benefits might be subsidized; if the subsidy is set equal to the external benefits resulting from the activity, the result will be efficient.

Not all externalities require public-sector involvement. Since people can and do bargain and negotiate, when small numbers are involved, people can come to private mutually beneficial agreements that are efficient.

A third source of market failure is *imperfect structure*. When the power to control prices is concentrated in a few hands, there is an incentive for overpricing and under production. Chapter 12 uses the polar case of monopoly to illustrate the social costs that can result from concentrated economic power. Monopoly firms earn higher profits by charging a price above what would prevail under competition. Consumers lose more than monopolists gain.

Redistribution of Income

The conclusion that the allocation of resources that would result from a perfectly competitive market system would be efficient does not ensure that it would be fair or equitable. In a market system, goods and services are rationed on the basis of willingness and ability to pay. The key is that those who value goods most end up with them; but the amount of value I am willing to give up to acquire a good depends on how much income I have.

Thus, only if the initial distribution of income and wealth is fair will the final allocation of resources and output be fair. There is a great deal of disagreement about what constitutes an equitable distribution. Every society provides certain basic goods to all regardless of income, even to those at the very bottom of the scale. The length of the list of basic goods, however, depends on the society's definition of justice.

The Impact of Taxes

With the exception of a tax on pure land rent, all taxes have impacts on decisions made by firms and households. When the behavior of private decision makers is affected by a tax, the allocation of resources changes. Some are better off, and some are worse off.

There are two very common fallacies in conventional thinking about taxes. First, many believe that if a tax is levied on an institution such as a corporation somehow, the burden of the tax disappears. Ultimately, however, the burden of all taxes comes to rest on individuals.

Second, many think that tax burdens fall on those on whom the law levies the tax. The discussion of tax incidence in Chapter 10, however, shows that tax burdens are shifted in a variety of ways.

Households bear tax burdens in two way. On the "sources" side of the income equation, households may find that the imposition of a tax causes wages, rents, or profits to rise or fall. On the "uses" side of the income equation, the prices of products may change. I am no better off if my income falls by 50 percent and prices remain the same than I would be if the prices of the things I buy double and my income remains the same.

It is important for tax policy makers to carefully analyze the incidence of taxes and of tax changes. Often even secondary effects can turn out to be large. Consider for example the debate in the United States over tax reform in 1985. One of the major objectives of the original reform package was to reduce the extent to which the current tax laws distort economic decisions.

There is no question that the current set of laws in the United States has led to a great deal of investment in low-income housing that otherwise would not have been undertaken. There is no question that moving to a realistic set of depreciation rules and reducing tax expenditures on housing would ultimately reduce the supply of low-income housing and drive up market rents. The Treasury Department's analysis of the tax reform proposals shows that low-income families will receive slight tax cuts and that they will presumably be better off. But low-income families spend between 35 and 50 percent of their incomes on rent, and a 10 percent increase in rent would swamp any modest tax reduction that they might receive.

Reducing distortions in the income tax will improve on the overall allocation of resources in society, but we should be aware of all the consequences. Without careful economic analysis, it is often difficult to identify who will gain and who will lose.

Taxes can affect any number of decisions that firms make including choice of technology and choice of location. Again, these impacts must be taken into account when alternative tax policies are being considered.

Other Aspects of Tax Policy Analysis: Growth and Excess Burden

Careful economic analysis shows clearly that the burden of most taxes exceeds the revenue collected. Theory suggests that the size of this *ex-*

cess burden depends on how much of a change in behavior results from the imposition of the tax. In general, partial taxes distort less than broad-based taxes. An excise tax on a single commodity certainly distorts decisions more than a general sales tax, since it might be avoided by simply buying a substitute.

Similarly, a tax on a commodity with an inelastic demand distorts less and imposes a smaller burden than a tax on a commodity with an elastic demand. A tax on pure land rent imposed on all uses of land at the same rate will distort no decisions and impose no excess burden at all, since total land is in inelastic supply and since a uniform rate will not change the relative attractiveness of various uses.

There is a great debate in the economic literature about the extent to which taxes affect the rate of savings and the supply of labor. The *supply-side* school of economics, associated with the Reagan administration, believes that taxes have seriously eroded the incentive to work, save, and invest in the United States. While economists studying actual behavior are not in total agreement about the size of these effects, most of the evidence seems to suggest that they are small.

Final Words

This book introduces students to a discipline and a way of thinking. A basic knowledge of economics is essential to the rational formulation of tax policy. It provides a structure and a logic within which many critical questions can be addressed. To many questions, however, there are no right answers; there are merely tradeoffs and much ambiguity:

How much growth should be sacrificed to achieve distributional equity?
What is the best tax base?
Who really bears the burden of the corporate income tax?

Many will find this residual ambiguity unsettling. We all want answers. But there are no simple answers to complicated questions; indeed, there are often not simple answers to simple questions.

People react in a variety of ways to the ambiguity and hard questions in economics. Some fall back on simple explanations and offer easy solutions; this seems to be the most common reaction of politicians. Others claim that the absence of agreement implies that economics provides no insights at all but they usually offer no alternatives.

To still others, complexity and ambiguity are challenges. When policy makers stop thinking, stop asking hard questions, and fail to look beyond simple explanations, we are in trouble.

Index

About the Author

Karl E. Case is an associate professor of economics at Wellesley College. He is also a visiting scholar at the Federal Reserve Bank of Boston and a lecturer on economics and tax policy in the International Tax Program at Harvard Law School. He has been associated with the Lincoln Institute of Land Policy for six years. He has been a consultant to tax reform commissions in New York and Massachusetts as well as several other state and federal agencies. He has previously published three books and numerous articles on public finance and urban economics. He is Associate Editor of the *Journal of Economic Education* and is Senior Associate for Economics and Public Policy of the Council of State Policy and Planning Agencies, the research arm of the National Governors Association.