
John Lodewijks; Mehdi Monadjemi

# Microeconomic Theory and Contemporary Issues

bookboon.com
The eBook company

MEHDI MONADJEMI, JOHN LODEWIJKS

# MICROECONOMIC THEORY AND CONTEMPORARY ISSUES

Microeconomic Theory and Contemporary Issues
1st edition

ISBN 978-87-403-1535-6
Peer review by John Lodewijks, SP Jain School of Global Management.

# CONTENTS

| | | |
|---|---|---|
| | About the Authors | 6 |
| | Preface | 8 |
| | Introduction | 10 |
| 1 | Consumer Behaviour and the Demand Curve | 14 |
| 2 | Theory of Production | 31 |
| 3 | Short Run and the Long Run Cost Functions | 47 |
| 4 | Perfect Competition | 63 |

| | | |
|---|---|---|
| 5 | Monopoly | 84 |
| 6 | Monopolistic Competition and Oligopoly | 99 |
| 7 | Input Markets | 118 |
| 8 | Externality, Public Goods and Government Regulations | 132 |
| | Endnotes | 144 |

# ABOUT THE AUTHORS

**Dr Mehdi Monadjemi**

Mehdi completed a B.S. in Economics from Utah State University and a M.S. and a Ph.D in Economics from Southern Illinois University, Carbondale, Illinois, U.S.A. His extensive experience in the banking and finance sector includes positions as Executive Director, Bank Refah and Bank Omran, Tehran, Iran, Economist, First Wisconsin National Bank of Milwaukee, London and Economist, Research Department, Reserve Bank of Australia. After eight years as Associate Professor of Economics, School of Economics and Political Science, The National University of Iran, he spent a further 20 years as an academic economist at the University of New South Wales, Australia including the Associate Head of the School of Economics position. He has held Visiting Scholar positions at Columbia University, London School of Economics and Political Science, and the University of Kent, Canterbury, United Kingdom. Currently he is visiting fellow at the School of Economics, University of New South Wales.

**Dr John Lodewijks**

John completed a Bachelor of Economics from the University of Sydney, Master of Economics from the University of New England and a M.A and PhD in Economics from Duke University, USA. He spent 22 years as an academic economist at the University of New South Wales, Australia including the Head of Department position. Thereafter he was Head of the School of Economics and Finance at the University of Western Sydney for a further five years. He is now associated with the S P Jain School of Global Management.

# PREFACE

Students often say to us: 'I want to ensure that the field of study I choose will offer a lucrative career once I graduate'. Unfortunately, our response often does not satisfy them.

We tell them that the global environment and information technology is changing so fast that there are no guarantees that the career that looks lucrative today will still be there in three or four years' time when you graduate. There has been a massive increase in the global tertiary-enrolment ratio which measures the share of student-age population attending university. The ratio has more than doubled over the last twenty years. So there are far more university graduates out there to compete with. To stand out you may need more and higher degree requirements. Where a bachelor's degree at pass level used to be sufficient, now employers are looking for degrees with merit performances or Master's degrees. Employers are also looking at the quality of the institution you graduate from – is it ranked internationally in the top 100? Finally there is the race between education and technology as the digital revolution and robots and automation and offshoring leads to traditional jobs disappearing while creating a whole new set of job opportunities.

So we tell students not to get too focused on the present – one recent study concluded that 47% of jobs now available may be computerized in the next 10 to 20 years – but think of what the future may entail. If your forecasting skills are rusty then go with where your heart lies. Follow your passion. Try not to be pressured into a career path that you are really not interested in just to meet your family's expectations. You will perform far better following a career path that really interests you as motivation is such a key factor in success. But be prepared to switch careers and employers multiple times during your working life as this is the state of the world now and hence you need good generic skills that allow you to be flexible and adaptable to seize any opportunities that come up.

The popularity of economics as a major in a business or even arts degree, or as a stand-alone Economics degree, has ebbed and flowed. Some students do not see that it leads to a well-defined career path – such as an accountancy degree does. Others find it too technical and abstract and do not see any clear connection with the real world. What these students fail to perceive is that a training in economics provides a generic set of skills including critical thinking, analytical and quantitative skills and to recognize the unintended consequences of a particular policy or action. These skills are clearly recognized in business, government and in awards of Nobel Prizes.

We believe passionately in the value of a training in economics for the betterment of society. This book deals with microeconomics, and is our companion volume to our earlier book dealing with Macroeconomics. The chapters cover the basic micro theory of consumers and firms and provide an introduction to the role of government. At the end of each chapter there is an application to a contemporary issue or development in the discipline to give the treatment a contemporary focus. It allows the student to see the myriad of issues that economic understanding can contribute to. We hope you find this material of interest and that it may persuade you to further your studies in economics.

M.M. & J.L.
October 2016

# INTRODUCTION

It is customary to divide economic analysis into micro and macroeconomics. Prior to the Great Depression of 1930s, the importance of micro and not macroeconomics was emphasized by philosophers and economic scholars. The classical and neo-classical economists assumed that in the long run an economy moves towards full employment. Any divergence from full employment is temporary and will be corrected by flexibility of wages and prices. The Great Depression of 1930s showed that capitalist economies can remain in less than full employment for an extended period of time.

The objection to classical economists' approach to automatic full employment was initially advanced by John Maynard Keynes in 1936 in the *Theory of Money, Interest and Employment.* Keynes argued that government intervention is needed to create employment when economy is faced with falling aggregate demand and depressed activity.

The foundation of microeconomics is based on ideas developed by Alfred Marshall in 1890. The classical economists assumed that value and prices are determined by the cost of production. Marshall argued that prices of goods and services are determined by interaction of demand and supply. The supply is determined by cost of production and profit maximization whereas demand is based on consumers' preferences. Hereafter, demand and supply have become the essential part of microeconomic analysis.

Initially it was assumed that individuals derive utility from consumption of goods and services. The concept of marginal utility was introduced to represent additional utility derived from consuming one more unit. The law of demand which is an inverse relationship between price and quantity demanded was explained in terms of diminishing marginal utility. As consumer possess' larger quantities of a good, each additional unit provides less utility and hence he is willing to pay lower prices for successive extra units. This method was called the Cardinal approach to derivation of the demand curve. Although the Cardinal approach is logical, it is not practical. Utility is not measurable: it is not possible to express additional utility by using a number. However, it is possible to say that one prefers a bundle of goods to another bundle or is indifferent. This approach is called the Ordinal approach to derivation of the law of demand. By using indifference curves and budget line, the ordinal approach has become the standard approach to derivation of the demand curve. Indifference curves are useful tools but quiet abstract and impractical. It is possible to drive the demand curve by using only budget lines. This is the approach which will be developed in this book.

Consumer behaviour and derivation of the demand curve based on budget lines are developed in chapter 1. Various measures of elasticity of demand are included in this chapter. Some discussion regarding statistical measurement of the demand curve is also offered. For the purpose of comparison, the derivation of the demand curve using indifference curves and budget line is discussed in the appendix of chapter 1.

In Chapters 2 and 3 the theoretical analysis of production and cost are developed for the derivation of the supply curve, based on firms attempting to maximize profit. Profit is measured by the difference between revenue and total cost of production. The revenue is price times quantity sold. The derivation of the cost functions is based on the theory of production and the unit price of resources. In these two chapters the law of diminishing marginal product is behind the shapes of the production and the cost curves.

Chapters 4 to 6 discuss output determination under four market conditions; perfect competition, monopolistic competition, oligopoly and monopoly. A perfectly competitive market is an ideal situation that does not exist in reality. However, knowing characteristic of this market allows one to measure the degree of competition in real world markets. Pure monopoly is another extreme case which is nonexistent. Most of the real world market situations fall into monopolistic competition and oligopoly. Several examples of market structures will be discussed in chapters 4 to 6. In addition, game theory as an example of oligopolistic behaviour is analysed in chapter 6. Chapter 1 to 6 are designed for price and output determination in product markets.

In chapter 7 prices of factors of production under perfect competition and imperfect competition are determined. The effects of labour unions on wages and employment and prevention of discriminatory wage payment is examined

Economic efficiency, public goods, externalities and the role of the government in provision of public goods and control of external diseconomies are discussed in chapters 8.

**Why is it important to study economics?**

Economics is the foundation discipline for studies in business, finance, accounting and related fields. Acknowledged as the queen of the social sciences, the scientific community recognizes the value of the discipline with the award of Nobel Prizes. Nevertheless, the study of economics is under threat. In the high schools, fewer and fewer students are taking economics in favour of more narrowly focused business studies. In the universities, Bachelor of Economics degrees are being jettisoned, and Departments closed in many institutions, and even there some are being turned out of Business faculties and into Arts.

Academically weak students are voting with their feet when faced with the option of choosing between challenging, analytical courses or descriptive, narrow courses with little intellectual challenge. That is the nation's loss. John Quiggin said it best and I paraphrase – A tertiary institution that offers a business degree without economics is not a real university, any more than one that that fails to offer Arts or Science degrees. Offering students supposedly vocational degrees with no real disciplinary foundation is a cruel fraud and a dereliction of educational duty. A cornerstone of a democratic society is economic literacy. Some of the toughest issues confronting society are rapid structural, technological and demographic change, macroeconomic turbulence, global warming, and growing income inequality; therefore some of our biggest challenges will be to design tax, health, welfare and education systems, along with competition and industrial laws that balance equity and efficiency, and robust macroeconomic policies for an integrated world economy. The performance of our economy impacts on the availability of jobs, mortgage interest rates, the value of the currency, the prices we pay in the supermarket and our overall standard of living. As such, economic performance affects everyone. We desperately need economically literate citizens that understand these public policy challenges.

Students should learn and graduate with the capacity to participate actively and responsibly in a diverse and changing world. It is the contestable nature of the economics discipline that makes it such a fascinating area of study. For students who care about how the world works, economics should be one of the most relevant and exciting courses to study. Yet many students view economics as abstract theory and fail to understand how it relates to actual decisions made by firms, consumers and governments. Partly it is our fault. We should focus more on the practical application of economic models, develop more fully basic data analysis skills and communicate our ideas better to non-economists and those students not pursuing economics majors. Our curricula may be becoming too narrow and front-loaded with technique at the expense of the generic skills so valued by employers related to analytical reasoning skills and critical thinking. An understanding of economic history and the history of economic ideas may temper a tendency to arrogance and make us more open to alternative approaches. Yet despite our limitations, the pendulum has swung far beyond what is healthy and essential for the well-being of our economy and society. The continuing decline of economics education has profound consequences. It should be reversed.

This book makes a modest contribution to understanding microeconomics. We also have a companion volume on macroeconomics and we encourage readers to sample that too:

Mehdi Monadjemi & John Lodewijks, *Money and Monetary Policy in an Open Economy* (Bookboon, London, 2015).

# 1 CONSUMER BEHAVIOUR AND THE DEMAND CURVE

Demand and supply have become an integral part of microeconomics since their introduction by Alfred Marshall in 1890. Demand and supply curves are essentially used for determination of price and output in a competitive market. However, they are also employed in different applications of the price system, resource allocations and welfare analysis. It is the bread and butter tool for the applied microeconomist.

In this chapter the demand curve is derived based on the cardinal approach to utility. Subsequently the derivation of the demand curve is presented using budget lines. Finally, various measures of the elasticity of demand and some discussions regarding estimation of the demand curve are also presented in this chapter along with some new approaches associated with behavioural economics. The ordinal approach to derivation of the demand curve using indifference curves and budget lines, based on utility maximization, is discussed in the appendix of the chapter.

**Utility and the Law of Demand**

This section discusses derivation of the law of demand using the cardinal approach. There are several assumptions for derivation of the demand curve:

a) Utility is the satisfaction derived from the consumption of goods and services.
b) Utility is measurable, the consumer is able to assign a specific number to show their satisfaction derived from consumption.
c) Consumers derive positive utility from consumption of an additional unit of a good.
d) Consumption of goods and services is subject to diminishing marginal utility (DMU).
e) Consumers are rational and, subject to their budget, choose goods and services such that their total utility is maximized.

Marginal utility (MU) is defined as additional utility obtained from consumption of an additional unit of a good. A hypothetical example of quantities of good X, their corresponding MUs and different prices willing to pay are presented in Table 1.

| $Q_x$ | $T_x$ | $MU_x$ | $\$P_x$ |
|---|---|---|---|
| 1 | 10 | 10 | 15 |
| 2 | 17 | 7 | 12 |
| 3 | 23 | 6 | 10 |
| 4 | 28 | 5 | 8 |

**Table 1** Derivation of the Demand Curve

In Table 1 $Q_x, T_x, MU_x$, and $P_x$ are quantity of X, total utility, marginal utility and the price of good X respectively. Marginal utility is: $MU_x = \frac{\Delta T_x}{\Delta Q_x}$. It is assumed that MU is diminishing for successive consumption of good X and hence the consumer is willing to pay a lower price for an additional unit of X. Information in table 1 is sufficient to establish an inverse relation between quantities of X and the price of X (the law of demand).

The demand curve is plotted in Figure 1. Changes in the price of commodity X cause quantity demanded to change in the opposite direction given that other factors remain constant.

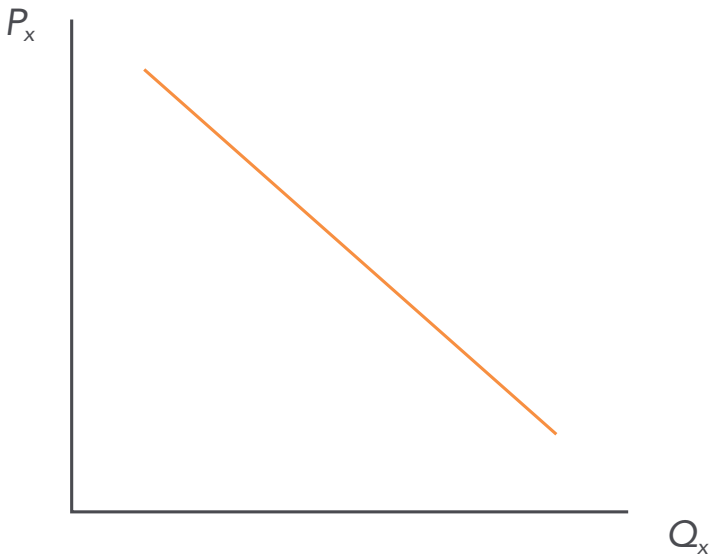


**Figure 1** Demand Curve

The assumption of consumers maximizing their total utility implies that every point on the demand curve represents the best choice for the consumer, subject to their income, price of good X and the price of other goods. The assumption of maximizing total utility means that marginal utility per dollar worth of all of the goods that are consumed must be equal. Assuming that consumer consumes two goods, X and Y, the condition for utility maximization is:

$$MU_x/P_x = MU_y/P_y$$

If this condition doesn't hold, the maximising consumer changes their desired combination of two goods such that again the above condition is satisfied. Suppose $MU_x/P_x > MU_y/P_y$, in this situation the consumer is better off to consume more of good X and less of Y. As he substitute X for Y, $MU_x$ falls (the law of diminishing marginal utility) and $MU_y$ rises. This process continues until the equality is restored again.

## An Alternative approach to Derivation of the Demand Curve

### The Budget Line

The assumption of measurable utility and consumer rationality are not realistic. Nobody is able to say how many units of satisfaction they derive from consumption of goods and services. In addition, consumer rationality requires knowledge of the availability of the entire set of alternative goods and their corresponding prices which is rather impossible. However, it is possible to derive the demand curve using an alternative approach without restrictive impractical assumptions.

Assume that the consumer has M dollars budget and wishes to spend it on consumption of Goods X and Y with prices $P_x$ and $P_y$ respectively. The allocation of his budget between expenditure on goods X and Y is:

$$M = XP_x + YP_y \quad (1)$$

Equation 1 is the consumer's budget line. It shows all combination of two goods that consumers are able to purchase when they spend their entire budget on goods X and Y. The consumer budget line is presented in Figure 2. Solving equation 1 for Y yields:

$$Y = M/P_y - \frac{P_x}{P_y}X \quad (2)$$

In equation 2, $\frac{M}{P_x}$ is the y intercept of the budget line, $\frac{M}{P_x}$ is the X intercept and $P_x/P_y$ is the slope of the line. The position of the budget line depend on M, $P_x$ and $P_y$. A change in his budget will move the budget line parallel up (an increase in M) or down (a decrease in M). When $P_x$ declines, M *and* $P_y$ remaining constant, $\frac{M}{P_x}$ moves to the right indicating that more good X can be purchased if the entire budget is spent on X.

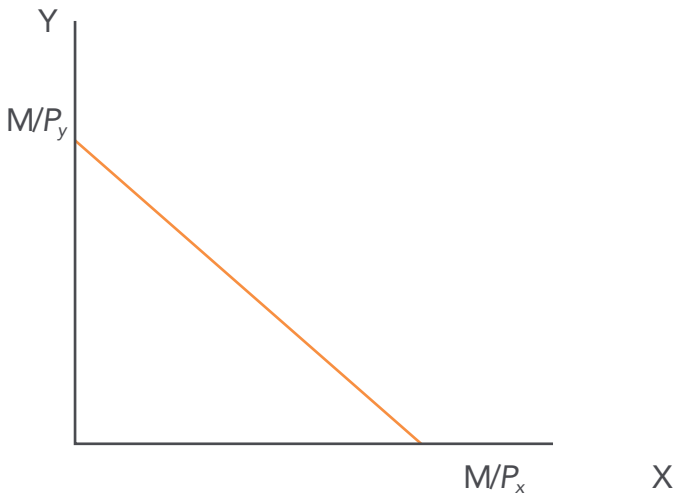


**Figure 2** The Budget Line

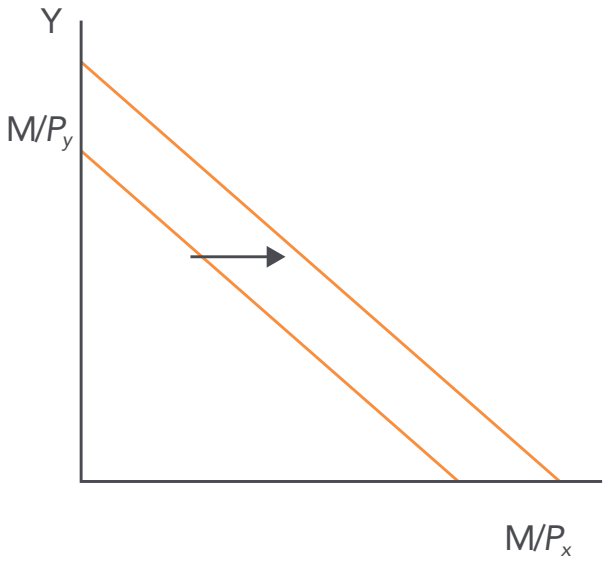


**Figure 3** An Increase in Budget

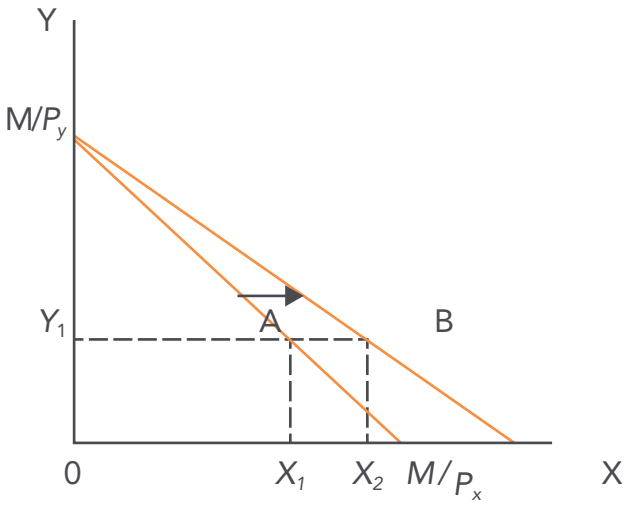


**Figure 4** A fall in Price of X

In Figures 3 and 4 it is assumed that good X is a normal good, consumers buy more when their purchasing power rises.

**Derivation of the demand Curve**

The law of demand can be derived using Figure 4. Initially at point A the consumer is purchasing $OX_1$ of good X and $OY_1$ of good Y. A fall in the price X, holding M and $P_y$ constant, causes the budget line to rotate to the right. The consumer moves to point B where he buys the same quantity of good Y and more of good X equal to $OX_2$. This is the law of demand. As price of X falls, consumer's purchasing power rises enabling him to buy a larger quantity of X. Here it is assumed that the entire rise in purchasing power is used to buy more of good X. However, it is possible to drive the same conclusion if the consumer decides to use his extra purchasing power to buy more quantities of both goods.

The demand curve can be derived after successive reductions of $P_x$ and larger quantities of X. Plot of all of the prices of X and their corresponding quantities yield an inverse relationship between price and quantity demanded which is the demand curve for good X.

The market demand curve is the horizontal summation of individual's demand curves. The market demand curve is negatively sloped provided that all of the individual demand curves are negatively sloped. The market demand curve shifts as a result of change in: consumer's budget, prices of substitutes or complimentary goods and consumers' taste for good X.

## Elasticity of Demand

The price elasticity of demand is defined as percentage change in quantity demanded in response to one percent change in price. The knowledge of elasticity is useful because it helps sellers to estimate the effects of price changes on their revenue. Also it can help labour unions to predict the effect of wage rise on total labour force earnings. Using the definition of elasticity, the price elasticity of demand for good X is:

$$e_{px} = \%\Delta Q_x / \%\Delta P_x = \Delta Q_x / Q_x / \Delta P_x / P_x = \frac{\Delta Q_x}{\Delta P_x} \text{ X } P_x / Q_x \qquad (3)$$

In equation 3 $Q_x$ and $P_x$ are quantity demanded and price of X respectively. The price elasticity of demand is negative because changes in price and quantity demanded move in the opposite direction.

In practice the absolute value of elasticity is used to show how elastic is the demand curve at a certain price.

The demand curve at a certain price is elastic when $e_{px} > 1$, inelastic when $e_{px} < 1$ and unitary elasticity when $e_{px} = 1$. Sometimes it is customary to refer to a demand curve being elastic or inelastic. This kind of observation is not strictly correct because the price elasticity varies along the demand curve from 0 to infinity. This point is illustrated in Figure 5.

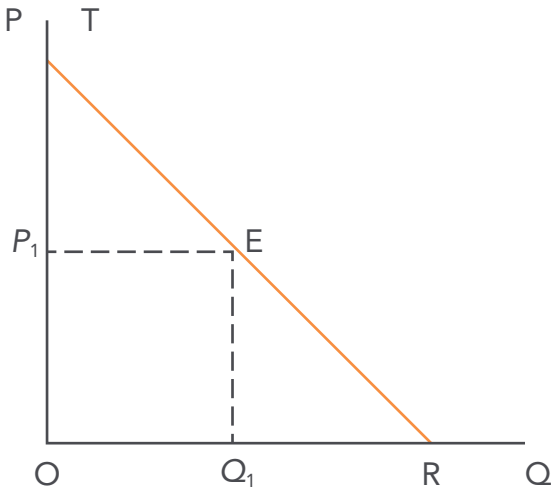


**Figure 5** Point Elasticity of demand

In Figure 5 a linear demand curve TR is plotted and price and quantity are measured on vertical and horizontal axis respectively. At point E the price and quantity demanded are $OP_1$ and $OQ_1$. The elasticity formula $\Delta Q_x/\Delta P_x \text{X } P_x/Q_x$ is called the point elasticity because it is measured at any point on the demand curve. At point E the elasticity is the inverse of the slope multiplied by the ratio of price to quantity at that point. Accordingly:

$$e_{px} = P_1\text{E}/\ P_1T \text{ X } OP_1/\ OQ_1 = OP_1/\ P_1T \text{ Since } OQ_1 = PE_1 \tag{4}$$

Depending on where the point E is located, the point elasticity could be greater, equal or less than unity. For a linear demand curve, $\Delta Q_x/\Delta P_x$ remains unchanged but the ratio of price to quantity varies along the demand cure. In the upper parts of TR elasticity is high because price is high and quantity is low. Towards the lower parts of the demand curve where price is low and quantity is high, the elasticity is less than unity. At E point elasticity is unity if E is located in the middle of the demand curve TR. In three cases of vertical, horizontal and hyperbola demand curves, the price elasticity of demand is uniform equal to 0, ∞ and -1 respectively,

The point elasticity of demand measures the elasticity when price and quantity changes are infinitesimal. The arc elasticity of demand is used when changes in price and quantity are large. The arc elasticity is the average of elasticity between two points on the demand curve. The arc elasticity formula is:

$$e_{px} = \frac{\Delta Q_x}{\Delta P_x}\ X\ (P_1 +\ P_2)/2/\ (Q_1 +\ Q_2)/2 = \frac{\Delta Q_x}{\Delta P_x}\ X\ (P_1 +\ P_2)\ /\ (Q_1 +\ Q_2) \tag{5}$$

Other measures of elasticity of demand are income elasticity and cross elasticity. The income elasticity of demand is percentage change in demand for good X as a result of one percent change in consumer's budget.

$$e_{Mx} = \frac{\Delta Q_x}{\Delta M} \text{ x } \frac{M}{Q_x} \tag{6}$$

In equation 6 M is the budget and $Q_x$ is the demand curve (a shift of the entire demand curve as opposed to change in quantity demanded which is movement along the curve).

$e_{Mx} > 0$ for normal goods and is negative for inferior goods.

The cross elasticity of demand is percentage change in demand for X as in response to one percent change in price of good Y.

$$e_{P_Yx} = \frac{\Delta Q_x}{\Delta P_Y} \text{ x } \frac{P_Y}{Q_x} \tag{7}$$

$e_{P_Y x} > 0$ for substitute goods (beef and pork) and is negative for complementary goods (tennis balls and tennis rackets).

**Elasticity and Total Revenue**

The knowledge of elasticity is important in estimating the effect of price changes on seller's revenue or the effect of wage rise on total earnings of the labour force. The total revenue is presented in equation 8.

$$TR = P \text{ X } Q \tag{8}$$

$$e_{px} = \frac{\% \Delta Q}{\% \Delta P} \tag{9}$$

In 9 if $\% \Delta Q > \% \Delta P$, $e_{px} > 1$. In this case when P rises Q falls by a larger amount causing TR in 9 to fall.

Also, if $\% \Delta Q < \% \Delta P$, $e_{px} < 1$, when P rises Q falls by a smaller amount causing TR to rise.

And finally, if $\% \Delta Q = \% \Delta P$, $e_{px} = 1$, when P rises Q falls by a the same amount causing TR to remain constant.

Accordingly, a seller who intends to increase his price needs an inelastic demand curve. In case of a decrease in price, higher total revenue requires an elastic demand curve (please verify this using the above explanations).

These analyses also apply to labour unions attempting to increase wage rates. Assuming that the demand curve for labour is downward sloping, an increase in wage rate requires an inelastic labour demand cure. Otherwise, the total earning of the labour force declines.

**Estimation of the Demand Curve**

The estimation of the demand curve looks simple but there are statistical complications.

One can use time series data on quantity and price of a good and statistically estimate the demand curve. However, the estimated relation is equilibrium points of demand and supply at each price which has been recorded, perhaps as a result of shifting supply curve. In other word, observed quantities are equilibrium points which could be a demand or a supply curve.

Figure 6 illustrates the above problem.

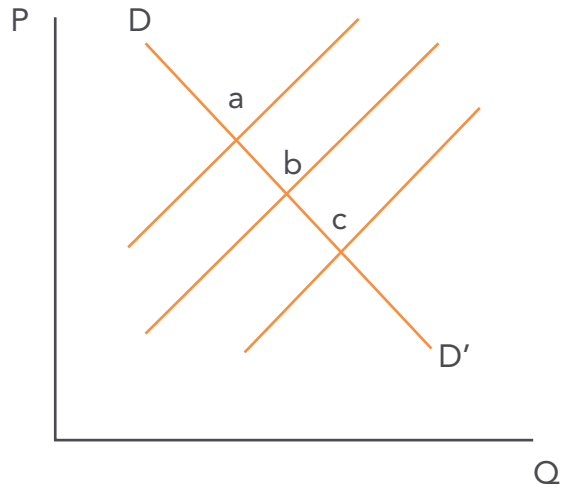


**Figure 6** Estimation of the Demand Curve

A researcher wishes to estimate the demand curve DD'. However, observations a, b and c are equilibrium points at the intersections of demand and supply. To avoid the problem researches have attempted to estimate the demand curve by including relevant determinants of demand as explanatory variables in the regression. By doing so, the estimated results may trace the demand curve such as DD' from the equilibrium observations a, b and c.

**Recent developments in demand theory**

The analysis covered so far is crucially dependent on key assumptions about consumer behaviour. Some of these assumptions have been relaxed over time and particularly so with the rise of behavioural economics. Let us examine some of these new ideas. One idea that is not so recent is that consumption is conspicuous in that it serves to reaffirm the consumer's role as part of a group. Consumers buy what they do because of the groups in which they are part. This is part of broader 'economics of conventions' approach as conventions, including social beliefs and institutions, are seen as shaping and constraining individual behaviour. The intersubjectivists, as they are sometimes called, reject the notion that consumers have well-defined tastes unaffected by social interactions, culture, institutions, or the consumption choices or well-being of others. Advertising and peer group pressure clearly affect consumption choices so that intersubjectivity is the norm, not the exception. All demand is intersubjective. The purpose then is to form concepts of non-atomistic agents subtle enough to capture the ambiguity and complexity of the intersubjective agent who is socially, culturally and institutionally embedded but who in turn shapes these structures. Intersubjectivists have applied these ideas to financial markets influenced by herd behaviour (imitation) and contagion of sentiments leading to non-optimal outcomes. For example, we watch television not only for the intrinsic please we derive from it but because we know others will watch it and this shared experience enables conversations to flow and this is central to social life. So demand for a program depends on the number of other people consuming it. This facet may lead to market failure in the sense of less diversity in programs and more sensationalist programs being screened. There is then a case for public broadcasting.

These are just some of the interesting ways that agent interdependence is being examined, and the links between economics and sociology, explored. Institutional economists note that interactive and partially malleable agents are entwined in a web of institutions that form and shape individuals. Institutions, like money, language, and family, are ingrained in habits of thought and conduct and have a temporal and ontological primacy over intention and reason. However, to model agents with endogenously formed preference functions would be very complicated and perhaps intractable. Yet few would doubt that in interactions between persons, envy, imitation, advertising, and so on play an important part in forming the utility or preference function and thus should be taken into account. The rise of experimental economics and behavioural economics shows psychological experimentation is being taken very seriously.

The basic problem is that of treating individuals as acting independently of each other. Game theory has played a key role of introducing intersubjectivity into economics. A model that assumes maximizing behaviour, market equilibrium, and stable preferences, is decidedly over-simplified. Evolutionary game theory, behavioural economics and experimental economics has enriched our understanding of complex behaviour patterns. For example, the concept of 'reciprocal fairness' reflects a willingness of individuals to sacrifice resources for rewarding fair and punishing unfair behaviour. Experimental results indicate that the fraction of individuals exhibiting some degree of reciprocal fairness is almost never below 40 precent and sometimes well above 60 per cent, while only between 20 and 30 per cent of individuals behaved in a completely selfish manner. Individuals who were unconditionally altruistic are virtually non-existent. There was a high degree of cooperation in market games.

Actual human behaviour is complex and rich. No economist would disagree. Many new approaches that deviate from methodological individualism, independent agents, perfect competition and perfect information are being explored. In the end, all models are abstractions; but do they generate interesting testable implications and predictions? That is the key issue.

## Appendix to Chapter 1: Indifference Curves and Budget Lines

The cardinal approach to the derivation of the demand curve is logical but impractical. It is based on the unrealistic assumption of measurable utility. Subsequently economists argued that the consumer cannot assign a number to express their utility but they are able to express their preferences in choosing goods and services. A consumer is able to say whether they prefer a bundle of goods to another or are indifferent between two bundles. This approach leads to the development of indifference curves (IDC).

An IDC shows all combinations of two goods that provide the same utility for the consumer. In other words, the consumer is indifferent between all of the bundles located on an IDC. An IDC is shown in Figure 7. All combinations of two goods X and Y, such as points A or B, yield the same level of utility and hence the consumer is indifferent between them.

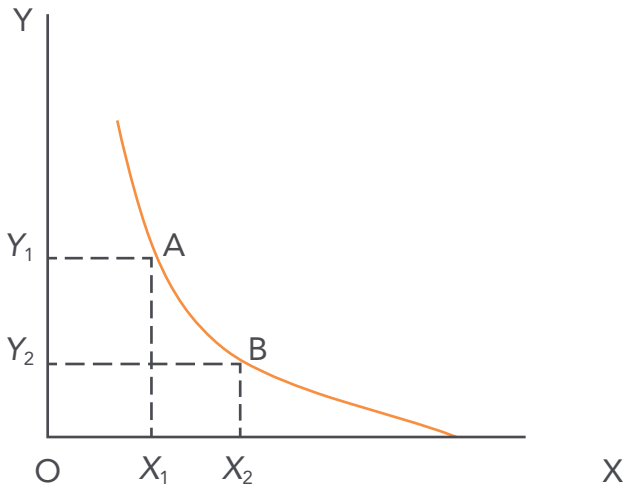


**Figure 7** an Indifference Curve

The ordinal approach assumes that more goods are preferred to less. This implies that a combination which includes more of both goods or has more of one good and the same amount of the other is preferred.

There are four important characteristics of IDC:

a. They are downward sloping. For total utility to remain constant on an IDC, increasing consumption of one good must be accompanied by a reduction of the other good, such as moving from A to B in Figure 7.
b. All of the points on higher IDCs located further away from the origin are preferred to the points on the lower IDCs.

c. IDCs are convex to the origin. As the consumer adds equal consumption good X, for his utility to remain constant, he must give up less and less of the other good. Each unit of X provides less and less satisfaction as more of X and less of Y is consumed. To remain on the same IDC (unchanged utility), the consumer is willing to sacrifice less and less quantity of Y for each additional unit of X.[1]
d. They do not intersect because if they do, every point on a higher IDC with more of one good and the same amount of other good cannot be preferred to all of the points on the lower IDC.

### The Consumer Equilibrium

The ordinal approach to derivation of the demand curve assumes that consumers are rational and choose a bundle of goods that maximizes their utility subject to their budget and prices of goods. In terms of IDCs, consumers attempt to reach the highest possible IDC allowed by their limited budget. Figure 8 shows the location of consumer equilibrium.

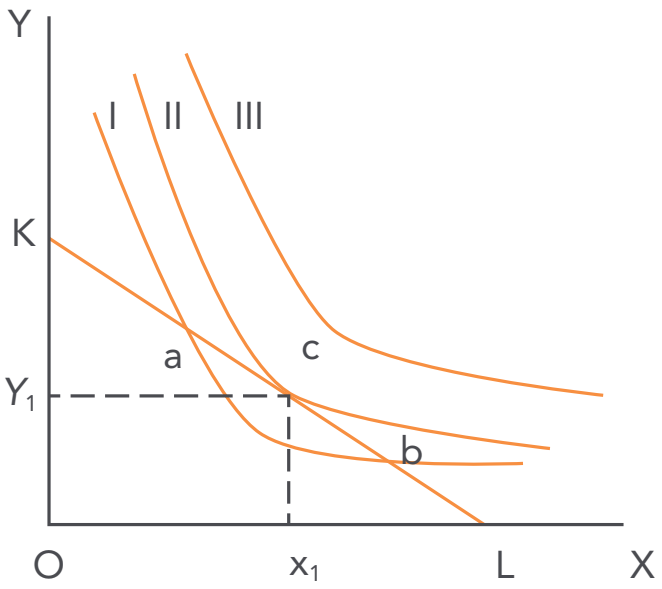


**Figure 8** Consumer Equilibrium

In Figure 8 three IDCs and the consumer budget line KL are presented. The consumer wishes to choose a bundle of two goods on highest IDC allowed by his budget line KL. The IDC III is unattainable. He has to stay on KL. Points a and b are inferior to c because c is on a higher IDC. Point c is the equilibrium because any movement to the right or left of c is on a lower IDC. His best choice is point c where he buys of $OX_1$ and $OY_1$ of good Y. At the equilibrium the consumer's budget line is tangent to the highest possible IDC. At the point of equilibrium the slope of IDC is equal to the slope of budget line, that is: $MU_x/MU_y = P_x/P_y$ or $MU_x/P_x = MU_y/P_y$.

**Changes in Consumer Equilibrium**

In Figure 9 an increase in consumer's budget shifts the budget line to K'L' which is tangent to IDC III at point d. At point d the consumer is buying more of both goods, assuming that both goods are normal.

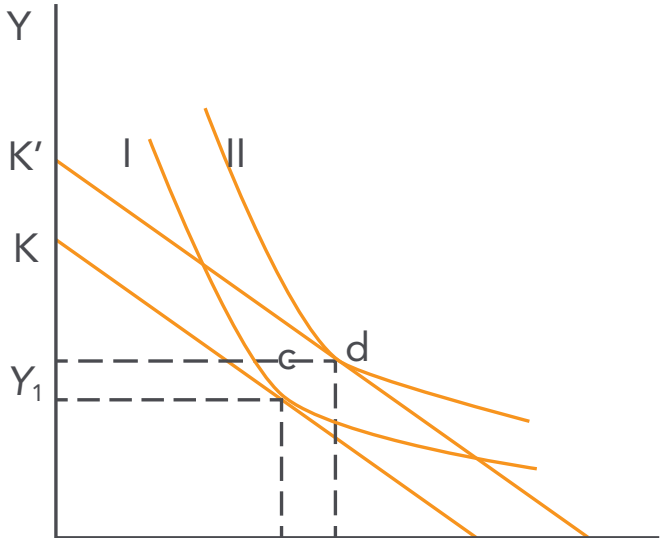


**Figure 9** an Increase in Income

### Income and Substitution Effect

When the price of X declines, because of income effect and substitution effect, the quantity of X demanded rises. The income effect is an increase in quantity demanded as a result of an increase in purchasing power of the consumer caused by a fall in price. The substitution effect is an increase in quantity demanded due to substitution of cheaper good for the other goods. The income effect and substitution effect of a fall in price of good X are presented in Figure 10.

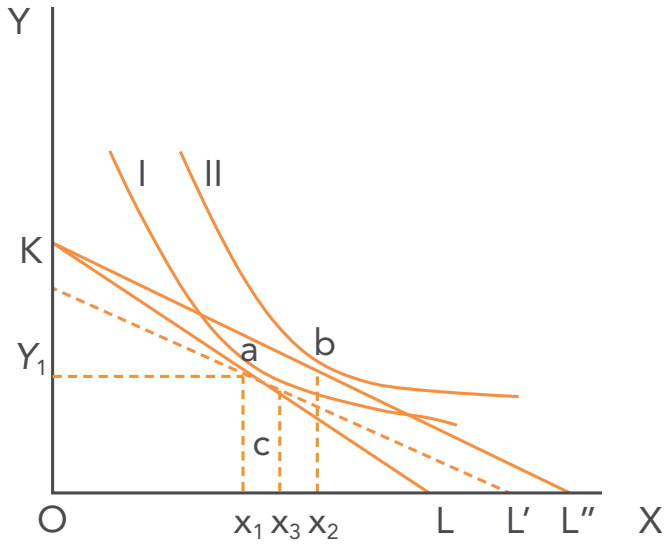


**Figure 10** Income and Substitution Effect

Initially the consumer equilibrium is at point a where he is consuming $Ox_1$ of X and $Oy_1$ of Y. As a result of a fall in price of X holding the budget and the price of Y constant, the budget line moves to KL' and the consumer equilibrium moves to point b. At b the desired quantity of X is $Ox_2$. The distance $x_1x_2$ is the total effect of a fall in the price of X which can be divided into income and substitution effect. The substitution effect can be isolated from the income effect holding purchasing power of consumer income (budget) constant. When the price X falls the purchasing power of consumer's income rises. This extra purchasing power can be offset by reducing consumer's money income. In Figure 10 holding purchasing power constant is shown by drawing a budget line such as L" parallel to Kl', lower price of X, such that it is tangent to the original IDC (implying constant purchasing power). The extra quantity demanded $x_1x_3$ is the substitution effect because it is as a result of falling price of X when consumer's purchasing power is held constant. If the reduction in money income is given back, the consumer moves from c to b. Hence, $x_3x_2$ is the income effect. The substitution effect is always movements along the IDC as consumer substitutes cheaper good for the other good. The income effect is movement from one IDC to another.

The substitution effect is always negative. The income effect is negative for normal goods and compensates the substitution effect. In case of normal goods the law of demand holds because the movement of price and quantity demanded are in the opposite direction. However, in case of inferior goods with a strong income effect the law of demand may not apply since the income effect offsets the substitution effect.

Using Figure 10, the demand curve can be derived by reducing the price of X further and tracing the effects on the quantity demanded. According to the cardinal approach every point on the demand curve shows the quantity demanded when the consumer is maximizing his utility subject to his income and the prices of goods.

# 2 THEORY OF PRODUCTION

Theories of production and cost are used for derivation of the supply curve. The cost of production depends on the employment of the factors of production. Hence, the cost of production is directly related to the employment of inputs for the purpose of production. Factors of production are labour, land, capital and entrepreneur. The theoretical discussion of production and cost is divided into short-run and long-run. The short run is the period of time that the quantity of some of the factors of production is fixed. In the long-run all of the factors of production are variable.

The short run theory of production assumes that there are two factors of production: a variable input, labour, and a fixed input, land. In the short run, gradual application of labour to a fixed size land leads to diminishing marginal productivity of labour. In the short run, the shapes of production functions and the cost functions are based on diminishing marginal productivity of the variable input.

The marginal product of labour is addition to total output as a result of an additional unit of labour.

$$MP_L = \Delta Q / \Delta L \quad (1)$$

Where $MP_L$, Q and L are marginal product of labour, total output and the quantity of labour respectively. Also a useful measure is the Average product of labour, $AP_L$.

$$AP_L = Q / L \quad (2)$$

Table 2.1 shows a hypothetical short–run production assuming that labour is the variable input and land is the fixed input.

| Land | Labour | Total Product TP | Marginal Product $MP_L = \Delta Q / \Delta L$ | Average Product $AP_L = Q / L$ |
|---|---|---|---|---|
| 1 | 1 | 10 | 10 | 10 |
| 1 | 2 | 15 | 5 | 7.5 |
| 1 | 3 | 22 | 7 | 3.1 |
| 1 | 4 | 26 | 4 | 6.5 |
| 1 | 5 | 26 | 0 | 5.2 |
| 1 | 6 | 24 | -2 | 4.0 |

**Table 2.1** a Hypothetical Short Run Production

In the above table the quantity of land is fixed at one unit.

In Table 2.1 AP reaches a maximum at 4 units of labour, MP is at maximum at 3 units of labour. When AP is rising, MP > AP. When AP is falling MP < AP. MP = AP when AP reaches a maximum in the neighbourhood of 3 and 4 units of labour. This relationship is true for any average and marginal. MP = 0 when the total product reaches a maximum and is negative when TP starts to fall.

A simplified graph showing the relationship between these cost concepts is presented in Figure 2.1.

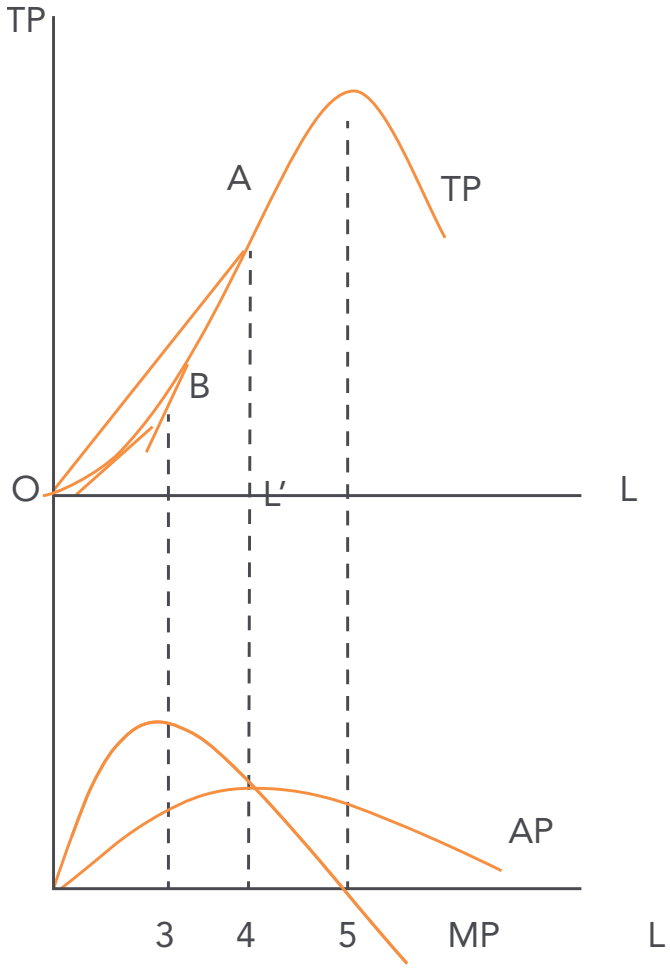


**Figure 2.1** Short Run Production Curves

The relationship between AP and MP can be explained geometrically. AP = TP/L = AL'/OL' in Figure 2.1. In general at any point on the TP curve, AP is the slope of the ray from the origin to that point, such as OA. AP reaches a maximum when the ray from the origin becomes tangent to the TP curve, point A on the graph.

MP = $\Delta Q$ / $\Delta L$ which is the slope of the TP, a tangent to the curve. The slope of TP increases as long as tangents lie behind the curve and declines when tangents are above the curve. MP reaches a maximum at the inflection point of TP when the tangent crosses the curve, point B on TP. Before B, MP increases as tangents are below the curve, and declines after B when tangents are located above TP. AP = MP and AP reaches a maximum at point A when the ray from the origin is tangent to TP.[2]

The short run production curves and their relationships are helpful in developing the short run cost functions.

## The Long Run Production

In the long run all factors of production are variable. A firm is free to choose any combination of resources. It is assumed that in the long run a firm chooses least cost combination of resources, that is, a combination that maximizes total output subject to a given cost. The optimum combination of inputs in the long run is determined using Isoquant and Isocost framework. Isoquants and Isocosts are similar to IDCs and budget lines except that they apply to producers' choice rather than consumers' choice.

An Isoquant shows all combinations of two inputs that produce the same output. An isoquant is presented in Figure 2.2.

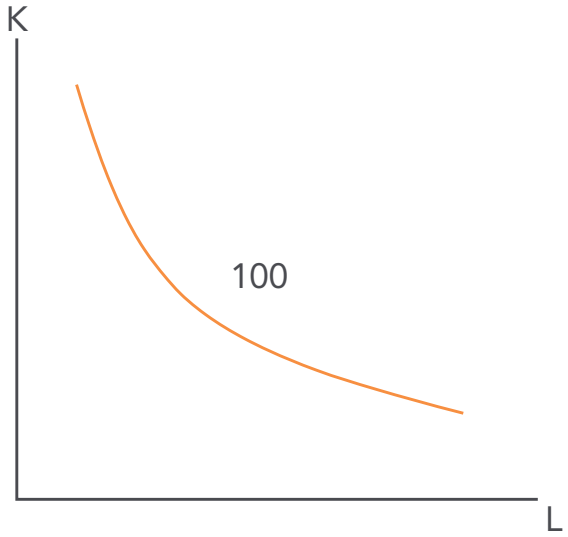


**Figure 2.2** an Isoquant

In Figure 2.2 K and L stand for capital and labour respectively. The isoquant in Figure 2.2 shows all combinations of K and L that can produce 100 units of output. Characteristics if Isoquants are similar to IDCs, i.e. they don't intersect, the higher curves represent more output, they are negatively sloped and are convex to the origin.

An isocost shows all combination of two inputs that a producer can purchase with a limited budget.

$$C = rK + wL \tag{3}$$

In equation 3 C, r and w are the outlay cost, unit price of capital and the wage rate respectively. Features of an isocost are similar to a budget line. An isocost is shown in Figure 2.3.

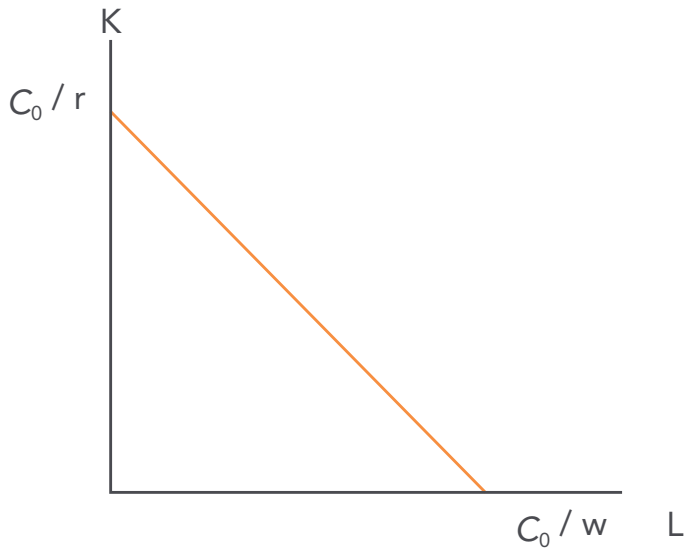


**Figure 2.3** an Isocost

Figure 2.3 shows all combinations of K and L that can be purchased with $\$ C_0$ budget. The intercepts of the budget line are $C_0$ / r and $C_0$ / w, showing respectively, the amounts of K and L that can be purchased if the entire budget is spent on one input.

### Least Cost combination of Resources

In the long run, a producer attempts to purchase a combination of two resources which maximizes output with a fixed budget, or produces a level of output that minimizes cost of purchasing two inputs.

The least cost combination is not attainable in the short run because the producer is not free to vary quantities of all inputs. The least cost combination is shown in Figure 2.4.

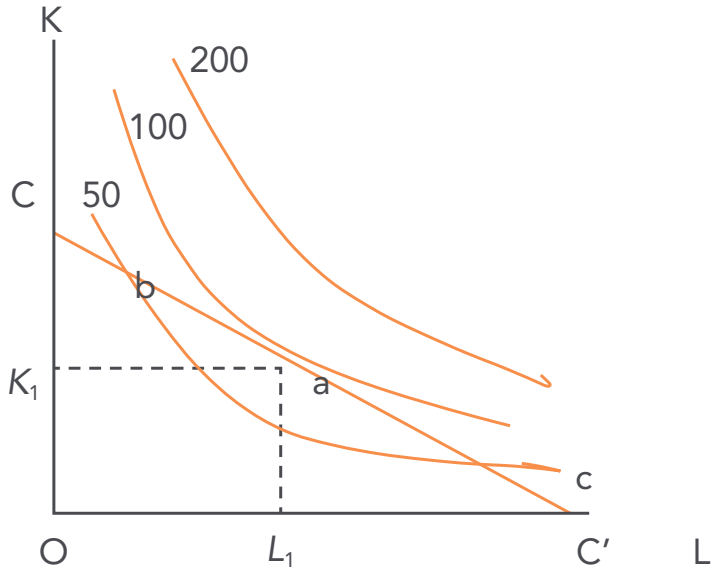


**Figure 2.4** Least Cost Combination

In Figure 2.4 there are three isoquants representing 50, 100 and 200 units of output. The CC' shows producer's limited budget that he wishes to spend on two inputs. The 200 units output is unattainable with his budget. He has to stay on his isocost. His best position is point a because any movement to the right or left of point a puts him on a lower isoquant or on a higher isocost. Accordingly, at point a by purchasing O $L_1$ labour and O $K_1$ of capital he maximizes output with a given cost or minimizes cost of a given level of output. At the optimum point the isocost is tangent to the highest possible isoquant. The slope of an isoquant is called marginal rate of technical substitution (MRTS) which is the ratio of two inputs marginal products. The slope of an isocost is the ratio of wage rate and the price of capital. At the point of tangency slope of the line and the curve are equal[3].

The condition for least cost combination of two inputs is given in equation 4.

$$\frac{MP_L}{MP_K} = \text{w} / \text{r} \tag{4}$$

The optimal least cost combination can be generalized to n inputs where the ratios of marginal product to price of an input are equal for all of the resources. This condition is useful for deriving the long run cost function.

**Returns to Scale**

Returns to scale applies to situations where inputs are changed by some proportion and output changes by the same or different proportion. Return to scale occurs in the long run when all of the inputs are variable. The returns to scale can be increasing, decreasing or constant.

Increasing returns to scale is a situation where a proportional increase in inputs causes a larger proportional increase in output. In case of decreasing returns to scale a proportional increase in inputs leads to a smaller proportional increase in output. Finally, constant returns to scale means that a proportional increase in inputs leads to the same proportion increase in output.

Using isoquants, constant, increasing and decreasing returns to scale for productions of 50, 100 and 150 (equal increases in output) units of output are shown in Figures 2.5, 2.6 and 2.7.

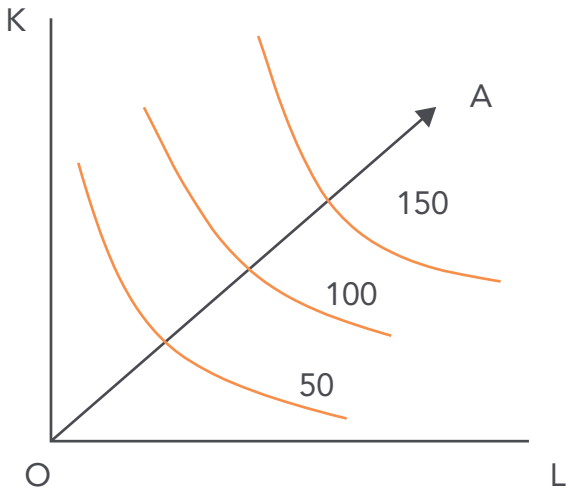


**Figure 2.5** Constant Returns to Scale

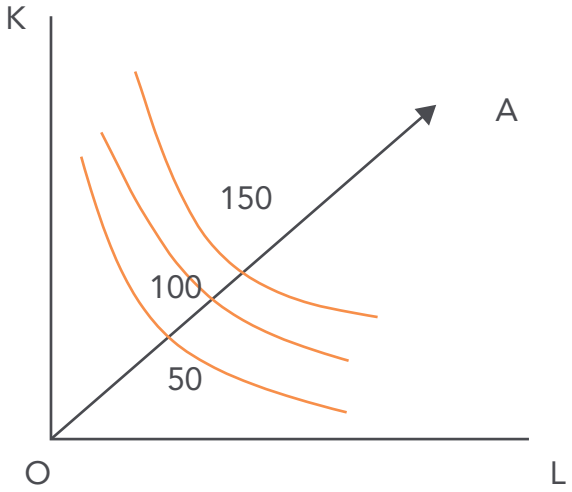


**Figure 2.6** Increasing Returns to Scale

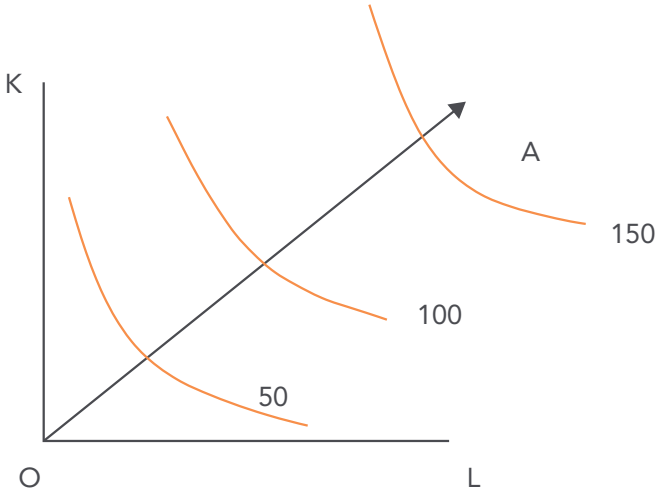


**Figure 2.7** Decreasing Returns to Scale

In Figure 2.5 on the ray OA isoquants are equally spaced. In case of increasing returns isoquants spaces on the ray OA gradually decline and finally in Figure 2.7 spaces between isoquants for equal increases in output become larger and larger.

Cobb Douglas (CDP) is a classic example of production function in economic literature. The CDP is presented in equation 5.

$$Q = AK^{\alpha} L^{\beta} \tag{5}$$

Where A, $\alpha$ and $\beta$ are the state of technology, capital share of output and labour share of output respectively. Using Cobb-Douglas production function, it can be shown that the degree of returns to scale can be determined by the sum of, $\alpha$ and $\beta$.

Multiplying K and L in equation 5 by a constant k yields:

$$A(kK)^{\alpha}(kL)^{\beta} = k^{\alpha+\beta} AK^{\alpha} L^{\beta} = k^{\alpha+\beta}Q \tag{6}$$

From 6 it can be deduced that the production function implies constant returns to scale if $\alpha + \beta = 1$, increasing returns to scale if $\alpha + \beta > 1$ and decreasing returns if $\alpha + \beta < 1$.

**Homogeneous Functions and Euler's Theorem**

A function is called homogeneous of degree n if each variable is multiplied by λ the function will be multiplied by $\lambda^n$.

It can be shown that Cobb Douglas production function is homogeneous of degree $\alpha + \beta$. If $\alpha + \beta = 1$, the Cobb Douglas production function is constant returns to scale. In this case the total output will be divided between two factors of production if each factor is paid its marginal product.

Equations 7 and 8 present marginal products of K and L respectively.

$$MP_k = \alpha A K^{\alpha-1} L^{\beta} \tag{7}$$

$$MP_L = \beta A K^{\alpha} L^{\beta-1} \tag{8}$$

$$K\alpha A K^{\alpha-1} L^{\beta} + L\,\beta A K^{\alpha} L^{\beta-1} = \alpha A K^{\alpha} L^{\beta} + \beta A K^{\alpha} L^{\beta} = \alpha + \beta(A K^{\alpha} L^{\beta}) = \alpha + \beta(Q) \tag{9}$$

Equation 9 is equal to total output if $\alpha + \beta = 1$.

**Estimation of Cobb Douglas Production Function**

Researchers have attempted to estimate the Cobb Douglas production function using least square regression. When the function is transformed into log linear, the coefficients of labour and capital become output elasticity of labour and capital.

Consider the Cobb Douglas production function in logarithm.

$$\log Q = \log A + \alpha \log K + \beta \log L \tag{10}$$

In equation 10 $\alpha$ and $\beta$ are output elasticity with respect to capital and labour. Also it is possible to test the validity of constant returns to scale by testing the restriction that $\alpha + \beta = 1$.

Felipe and Adam (2005) attempted to re-estimate the original Cobb Douglas production function initially estimated by Cobb and Douglas (1928) using US manufacturing data 1898–1922. The OLS regression results showed that, α = 0.75 for labour and β = 0.25 for capital. The authors imposed constant return to scale providing empirical support for the existence of the aggregate production function and the neo-classical marginal productivity theory of distribution.

Subsequent work in this area criticized the empirical results of Cobb-Douglas research arguing against existence of an aggregate production function. They maintained that the estimated function is the identity of share of profit and wage in national income. The authors re-estimated the Cobb-Douglas function using the original data. The results of estimation provided support for the constant return to scale when a time trend was not included in the regression. However, the coefficient of the capital stock became negative once a time trend was added to represent the effect of technical progress. They concluded that empirical results of an aggregate production function and return to scale are suspicious if one believes that the results yield identity of wage and profit in national income. There are other deeper methodological concerns about aggregate production functions relating to the measurement of capital that are beyond the scope of this introductory treatment.

**The Theory of Production and Long-Term Economic Growth**

The microeconomic issues that we have covered in this chapter relating to factors of production, productivity and the production function approach can all be related to the examination of economic growth. We have expectations that our standard of living and incomes will grow over time. We expect that we will have greater material possessions and disposable incomes than our parents and certainly our grandparents.

Yet the first point to note is that economic growth is historically speaking a relatively recent development. Angus Maddison, a renowned expert on the empirics of growth, tells us that over the past millennium, world population rose 22-fold. Per capita income increased 13-fold, world Gross Domestic Product (GDP) nearly 300-fold. This contrasts sharply with the preceding millennium, when world population grew by only a sixth, and there was no advance in per capita income. In other words, there was a period of one thousand years when the world experienced no real economic growth whatsoever. Really it has only been since 1820 that world growth rates have been more dynamic.

Per capita income growth is of course not the only indicator of welfare. To put this growth experience in perspective we might note that back in the year 1000, the average infant could expect to live about 24 years. A third would die in the first year of life, while hunger and epidemic diseases would ravage the survivors. Most of the improvements in these respects have occurred since 1820.

In relatively recent years, the period 1950 to 1973 stands out as a 'golden age' of unparalleled prosperity. The world economy grew faster during this period than it had ever done before. World per capita GDP rose nearly 3 per cent per year which implies a doubling in average income every 25 years. The overall momentum of growth has decelerated sharply since 1973. The resurgence of rapid growth over the last decades has been confined to only a few regions of the world and has generally not lasted long.

Let us explore some of these issues using some of the concepts already developed in this chapter. We define economic growth as increases in a country's ability to produce goods and services over time. One commonly used measure of a population's standard of living is real per-capita income. A growing population requires a growing real income to maintain its standard of living. Economists believe that rising real incomes are produced by the accumulation of capital and the rising productivity of labour. Labour productivity is affected by growth in physical and human capital, natural resources and technology. In order to view the relationship between output and the input of resources, economists use a production function, as mentioned earlier in this chapter. To see what happens to per capita output, we often use the per-worker production function. Using the per-worker production function, the law of diminishing returns to capital can reduce growth rates in an economy as output per capita increases.

Generally speaking, economic growth may be caused by:

- an increased availability of resources
- an improvement in the quality of resources
- technological improvements that make better use of resources.

Nobel Prize winning economist Simon Kuznets argued that changes in labour and capital accounted for only 10 per cent of the increase in economic growth in developed market economies. The other 90 per cent was due to improvements in the quality of inputs (that is, technological change). This points to the importance of research and development in stimulating economic growth, as improvements in technology arise from research. New capital embodies the fruits of scientific inquiry and serves as the primary engine for growth.

We now look at a case study of two countries to apply some of the ideas mentioned in this chapter. We examine the growth experiences in Japan and South Korea using the findings of a book written by Dirk Pilat titled 'The Economics of Rapid Growth: The Experiences of Japan and Korea' (Cheltenham: Edward Elgar, 1994).

Dirk Pilat has tried to analyse and quantify the main sources of post-war growth in Japan and South Korea and to measure the extent of the 'catch-up' to American levels of income and productivity. In both countries over the four decades since 1953, average income grew at a remarkable rate of almost five and a half per cent annually. Interestingly, after 1973, Japanese average income grew at less than half its earlier rate while the Korean rate almost doubled. For Japan, two-thirds of the measured economic growth arose from increased factor inputs (capital and labour). For Korea this figure was closer to 80 per cent.

The most important single factor for Japan was fixed private (non-residential) investment, which accounted for 32 per cent of the overall economic growth. Investment is the primary means of expanding the productive capacity of the economy and the way through which new technology can enter the production process. Capital investment in Japan has been financed primarily by domestic saving – foreign saving has played a very minor role. With respect to technology generation, Japanese research was aimed primarily at commercial application with little being spent on basic research.

Pilat notes that government in Japan has played an important role in directing investment to high growth industries and in gaining access to modern technology. Japan followed an industrial policy directed at the creation of long-term dynamic comparative advantage. In Korea, the government's performance is rated as 'outstanding'. It has allowed a broad range of industries to develop under a neutral trade regime by offsetting the discriminatory effects of import protection through various export subsidies. Hence, protection for domestic industries was matched by incentives given to exporters. Secondly, through the selective promotion of infant industries the government promoted activities with long-term dynamic benefits to the economy. However, to encourage efficiency, these infant industries were required to sell an increasing share of their output on world markets.

Japanese and Korean productivity performance, relative to US standards, is revealing. In all sectors, except agriculture, Japanese productivity levels were gaining on the US. A number of sectors have caught up with US productivity levels and some have surpassed this level. Over the 1975–85 period almost all Japanese manufacturing branches made significant productivity gains relative to the United States. Manufacturing is still the dominant sector in world trade, with a share of seventy per cent, and is important for Japan because productivity growth in this sector accounts for 35 per cent of the total catch up to US levels. Starting from a productivity level slightly above 20 per cent in 1953, Japanese manufacturing reached more than 90 per cent of the US level in 1990. For the economy as a whole the rise was from 20 per cent in 1953 to almost 65 per cent in 1990. By this later date capital equipment per person had become almost the same in Japan and the United States. Similarly, the difference in the quality of the labour force had become fairly small.

Korea also made substantial progress in terms of relative productivity, albeit from a far lower base. For example, agricultural productivity is less than 5 per cent of the US level. Output per person employed for the whole economy rose from 18 per cent of the US level in 1975 to 31 per cent in 1985. Over the period 1963–90, 42 per cent of Korea's total productivity growth can be attributed to the manufacturing sector. Yet in this sector capital per person employed was only 60 per cent of the US level in 1987, illustrating that there is still a big gap with the United States. On the other hand, in terms of education, the Korean workforce had reached a quality level of almost 90 per cent of the United States.

Pilat then turns his attention to international comparisons of productivity, with the analysis broadened to include more than ten countries, and competitive performance in terms of unit labour costs and trade policy (neither Japan or Korea have been very open to trade). It was found that for Japan those sectors with the highest protection from imports have the lowest productivity growth. Alternatively, sectors with low levels of protection have experienced the fastest productivity growth. Surprisingly for Korea the relation is found to be the exact reverse! Although for both countries there is a strong positive link between export growth and productivity performance.

The detailed quantitative research that Pilat undertakes is indispensable in our understanding of growth processes in these nations. It also shows how the fundamental microeconomic concepts of the theory of production introduced in this chapter can be used to illuminate discussion of vital issues affecting our material standard of living.

# 3 SHORT RUN AND THE LONG RUN COST FUNCTIONS

Similar to the theory of production, in the short run the total cost of production consist of fixed costs and variable. For example, the costs of purchasing land or building a factory are fixed costs because the quantity of these inputs cannot be changed quickly. On the other hand, labour, and raw materials are usually considered as variable costs. In the short run fixed costs are independent of the level of output whereas variable costs are function of the level of output. The short run total cost, the sum of variable cost and the fixed cost, are presented in equation 3.1.

$$STC = TFC + TVC\ (Q) \tag{3.1}$$

In 3.1 Q is the level of output. In the long run all costs are variable. The long run cost function is derived from the least cost combination of inputs.

**The Short Run Cost Functions**

The economic costs of production are different than accounting costs (outlays). The economic costs are explicit costs and implicit costs. Explicit costs are outlays that entrepreneurs have to meet in-order to produce. Implicit costs are opportunity costs of using the owner's own resources. This is a minimum that owner must make in-order to remain in business, sometimes it is called normal profit. Implicit costs are the same in the short run as well as in the long run. The explicit costs are different in the short run than the long run because there are no fixed costs in the long run.

The short run cost function can be derived from the short run production functions. Consider the following short run average and marginal cost curves:

$$AVC = TVC / Q = L \times w / Q = w\,(L/Q) = w\,(1 / AP_L) \tag{3.2}$$

Where SAVC, and w are short run average total cost and the wage rate respectively.

The wage rate is constant when labour is hired under perfect competition (this assumption will be discussed in the next chapter). In 3.2 SRVC is inversely related to $AP_L$ multiplied by a constant. This means that when $AP_L$ is rising, AVC falls, when the former falls the latter rises and finally, AVC is at minimum when $AP_L$ is at maximum.

Similarly it can be shown that $MC = w\,(1/MP_L)$

$$TVC = w\,L \qquad MC = w\,\Delta L$$

$$MC = w\left(\frac{\Delta L}{\Delta Q}\right) = w\left(\frac{1}{MP}\right) \tag{3.3}$$

Equation 3.3 shows that MC and MP are inversely related after multiplying by a constant.

Equation 3.3 can be used to derive the short run TVC and TC. AVC and MC and ATC can be derived using AP and MP in Table 1 in chapter 2.

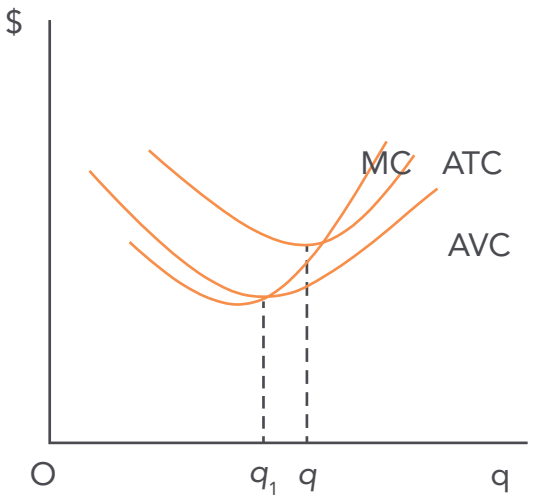


**Figure 3.1** Unit cost Curves

In Figure 3.1 the MC curve passes through the minimum points of AVC and ATC. When AVC is falling MC < AVC and when the former is rising, MC > AVC. The same relationship holds between ATC and MC. The difference between AVC and ATC is average fixed cost, AFC, which decline continuously but never reaches zero.

$$AFC = TFC / Q \tag{3.2}$$

In 3.2 TFC is fixed as q rises, causing AFC to fall continuously. The AFC is a rectangular hyperbola where the area under the curve remains fixed. The AFC curve is shown in Figure 3.2 where all of the rectangles under the curve, representing the TFC have the same area.

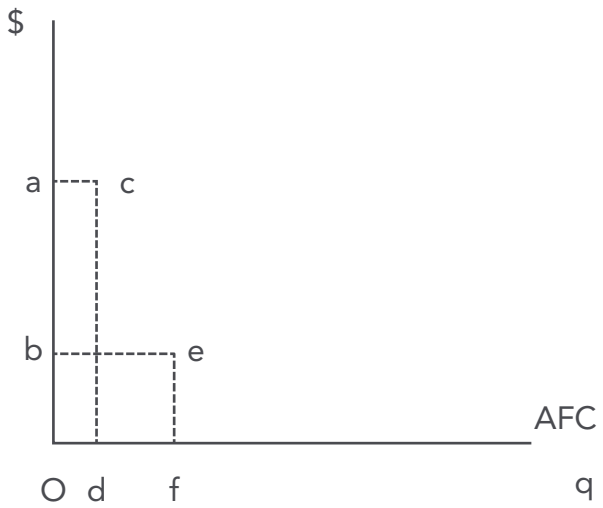


**Figure 3.2** the Average Fixed Cost Curve

In Figure 3.2 all of the rectangles such as acod and beof have the same areas equal to the total fixed cost.

The TC, TFC and TVC are plotted in Figure 3.3. The shape of TVC is based on falling and rising MC which is the slope of TVC curve. The vertical distance between TVC and TC curves is the TFC.

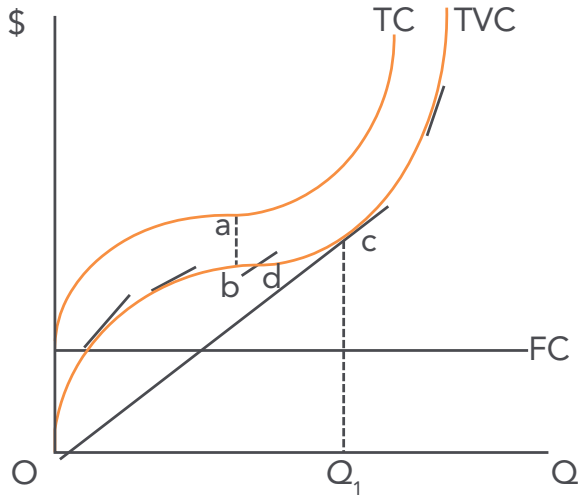


**Figure 3.3** the Total Cost Curves

In Figure 3.3 The difference between TC and TVC ab, is the AFC which remains fixed irrespective of the level of output. The slope of the TVC is MC (a tangent to the curve) which gradually declines till point c and then increases. At c AVC reaches a minimum where AVC = MC. At any point on the TVC curve, the AVC is the slope of the line connecting the origin to a point such as c (similar to AP curve). AVC gradually declines, reaches a minimum at point c and then rises. At c, AVC = MC. The MC curve reaches a minimum at the inflection point of TVC curve such as point d where the tangent cuts the TVC curve (the inflection point). Similar relationship holds between ATC and MC. The minimum of ATC is at a quantity larger than the minimum of AVC because when AVC reaches a minimum; AFC is still falling causing ATC to continue to fall.

### Geometry of the Short Run Cost Functions

Figures 3.4 shows how AVC and MC are related to the TVC.

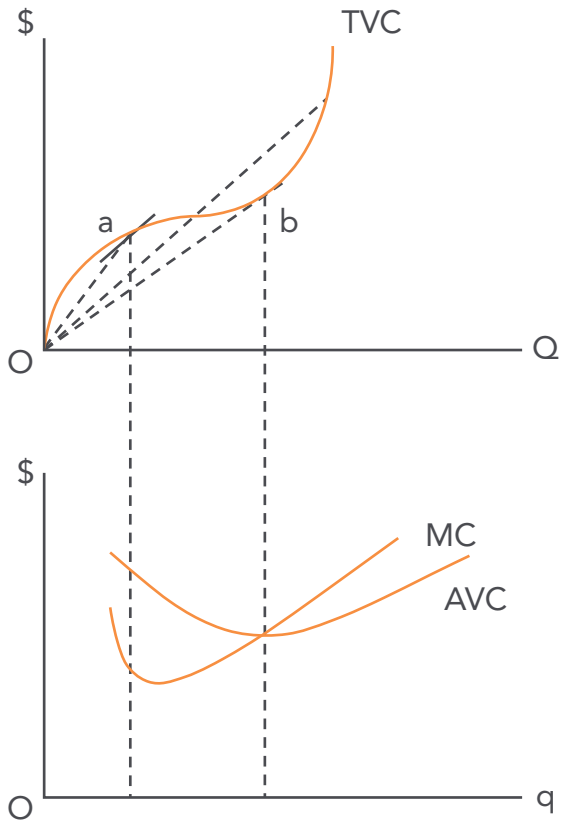


**Figure 3.4** Total Variable Cost

In Figure 3.4 AVC reaches a minimum where the ray from the origin such as ob is tangent to the TVC curve. At this point AVC is equal to MC. Before point b rays from the origin become flatter, after point b those rays gradually become steeper. At point a where the tangent to TVC curve cuts the curve, MC reaches a minimum. By comparing slopes of the TVC curve and slope of the rays from the origin, it can easily be shown that when AVC is falling MC < AVC, and when AVC is rising MC > AVC (this exercise is left for the reader).

### The Long Run Cost (LRC)

The long run is the planning period for the firm. There are no restriction on the scale of plant and the firm can choose any quantity of fixed input which minimizes the cost of producing a given level of output. This point is illustrated in Figure 3.5. Assume that the in the short run the level of fixed input is $F_1$ which is consistent with minimizing the cost of producing 100 units of output. The cost of producing 100 units is a point on the LRC curve. Other points on the LRC curve can be derived by varying the quantity of output with different scale of plants (fixed input) that minimizes the cost.

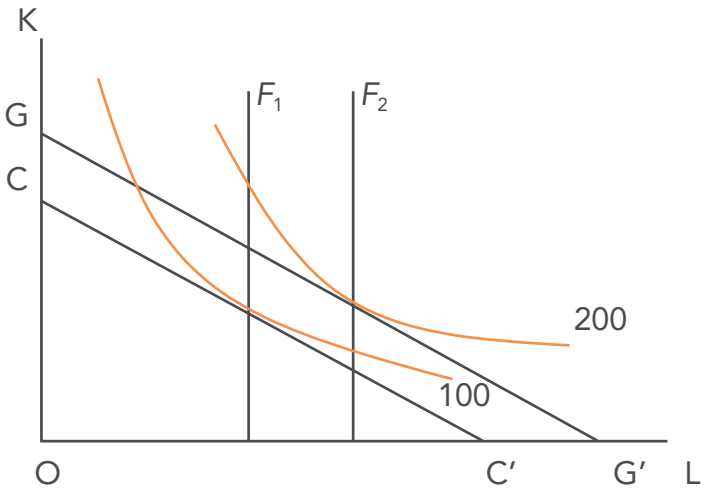


**Figure 3.5** Long Run Cost

In Figure 3.5 CC' for producing 100 units of output with a scale of plant $F_1$ and GG' for production of 200 units using $F_2$ as the scale of plant, are long run costs of production. Similarly other long run pairs of output and cost can be derived. Plotting these points yield the LRC curve. The precise shape of the LRC curve is unknown. However, the above theoretical discussions suggest that there is a positive relation between output and LRC.

Many economists argue that the average long run cost curve (ALRC) is also U shape and is made of tangency points of infinite number of short run ATC curves. This type of ALRC is explained by the economies and diseconomies of scale.

Consider Figure 3.6 where S1, S2 and S3 are three Short run ATC curves. To produce a given level of output such as $Oq_1$, because of fixed input, the firm must operate with S2. However in the long run can choose a scale of plant such as S1 which minimizes cost of producing $Oq_1$ Accordingly, point a is a point on the LRAC curve. Similar explanation can apply to point b. The LRAC curve (the envelope curve) shows the least average cost of producing different levels of output. The long run cost can never exceed the short run cost for a given level of output. If the short run cost is not the least cost (because of fixed input) in the long run the firm can always reduce the cost of producing that output by choosing different scale of plant (fixed input).

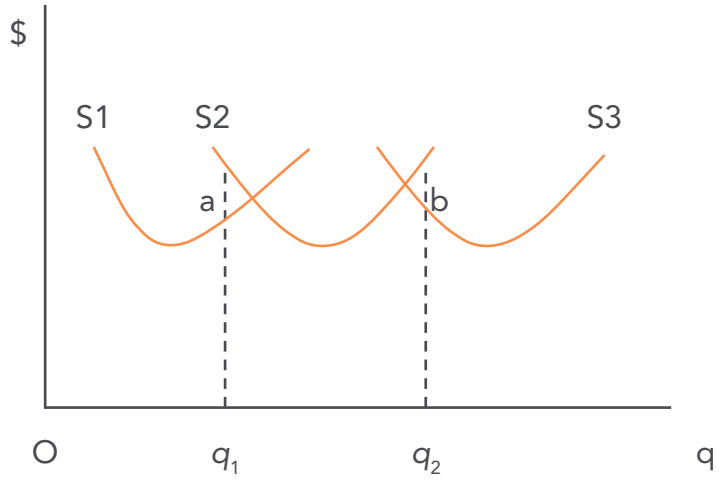


**Figure 3.6** the Long Run Average Cost

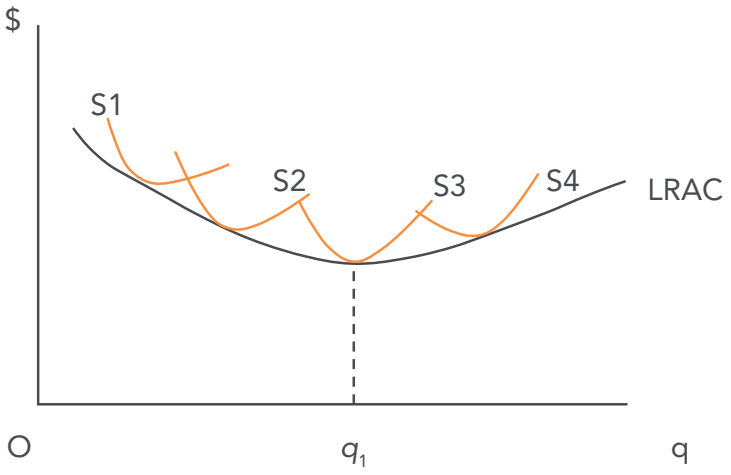


**Figure 3.7** the Envelope curve

In Figure 3.7 the LRAC curve (the envelope curve) is constructed from the infinite number of ATC curves where at the point of tangency the average cost of producing a given level of output is at minimum. This doesn't mean that the point tangency is the minimum point of ATCs, There is only one ATC such as S3 which its minimum is tangent to the minimum of LRAC curve.[4]

### The Shape of the LRAC Curve

The theoretical discussion of the long run cost provides no support for any specific shape of the long run cost function except a positive relationship between long run cost and output. However, A "U" shaped LRAC has been promoted based on economies and diseconomies of scale.

The falling section of the LRAC is based on the economies of scale causing the average cost of production to fall as the scale of operation expands and larger quantities are produced. The economies of scale is based on specialization of labour as the scale of operation expands, discounts on bulk purchase of raw material and lower per unit cost of installing larger than smaller equipment. All of these benefits are realized when the firm expands its scale of operation and produces larger quantities of output.

Gains from the economies of scale are not unlimited. At some stage diseconomies of scale more than offset the economies causing the LRAC starts to rise. One of the most important diseconomies of scale is the management of large enterprises. Inefficiencies arise when management control becomes weaker as the scale of operation expands. Decentralization of management responsibilities is costly and requires substantial cooperation and internal communication between different levels of management.

It is very difficult to determine at what quantity of output the LRAC stats to rise. Some argue that LRAC can remain at the minimum if when output expands if management controls are effective.

**Real World Applications**

The above discussion has been abstract and theoretical. Yet minimizing costs and economies of scale are fundamental to modern business. Let us illustrate this with a discussion of offshore outsourcing.

In 2006 the influential U.S Council on Foreign Relation's publication Foreign Affairs featured an article by Alan Blinder titled 'Offshoring: the Next Industrial Revolution?'. Offshoring refers to business processes that were originally performed 'in-house' being moved overseas. Blinder (2006:114) noted that 'We have so far barely seen the tip of the offshoring iceberg, the eventual dimensions of which may be staggering'. The next Industrial Revolution – the Information Age – revolves around advances in information technology that make many services now tradeable via the flow of packets of digitized information. Services that are easily delivered electronically can be moved offshore while those that continue to rely on personal face-to-face contact will remain non-tradeable. Within a firm this distinguishes front-office services involving direct customer contact that remains in-house with back-office services that may be performed – outsourced – in overseas locations.

Two years later, Youngdahl and Ramaswamy (2008: 213) commented that offshoring was "perhaps the most important phenomenon transforming the workplace". They highlighted India's technology parks where 250 engineers were developing IT applications for the Bank of America, finance staff were processing home loans for another American bank, radiologists were interpreting CT scans for an American hospital, and seven PhD qualified molecular biologists were undertaking research for a pharmaceutical company. Similarly examples could have been provided from Manila, Shanghai or Budapest. The offshoring movement had progressed far beyond call centres to increasing level of complexity in design, manufacture and service functions.

**Outsourcing and Offshoring: The basic principles**

Improvements in the quality and reductions in the price, of information technology, combined with the increased technical skills of those in relatively low-wage countries have facilitated the growth of outsourcing and offshoring. Firms are now continually reassessing what is core to their business, and so must remain in-house to maintain their competitive advantage, and what are auxiliary activities that can be contracted out or outsourced to firms or individuals locally or increasingly, overseas. The increasing division of labour, specialization or fragmentation of the supply-chain/value-chain is a clear feature of the modern business landscape. Firms are slicing up the firm's value chain into constituent parts and deciding how and where each part will be produced. 'Core' segments remain in-house while other segments may be contracted out to other firms or subsidiaries either nationally or internationally. Yet more and more parts of what used to be regarded as core activities are being outsourced as global competition has intensified. Initially outsourced activities were low-skilled but increasingly high-value functions such as R & D, design and engineering are being performed outside the firm. While **the main driver has been cost saving**, and this is still very important, so is access to knowledge and talented teams of people that are not available locally.

Each firm's function – R & D, Marketing, Finance, Production – can be sliced into hundreds of sub-activities. Each of these sub-activities can be performed internally in the firm or outsourced depending on the direct cost, including the transaction costs of organization, communication, coordination and quality control issues, associated with dealing with firms or small groups of specialists that are not directly related to the firm in question. In addition, there are sensitive issues relating to intellectual property rights that need to be protected from imitation. This is part of a broader strategic concern that the 'partner' firm used for out-sourcing may enhance their capabilities and knowledge in the process and later prove a competitive threat to the incumbent firm.

A longer-term consequence is that if a certain specialist skill or capability is outsourced for a period of time then that domestic capacity may shrink or be extinguished in the incumbent firm. For example, as more IT-programming is turned over to offshore providers, organizations may lose those in-house capabilities. This is akin to the notion that long term unemployed workers find that their human capital has depreciated and become obsolete the longer they are out of the work force. If a firm loses that capability it becomes dependent on the outsourcer or foreign firm and hence more susceptible and vulnerable to the threat of supply disruptions and price hikes.

The relationship between the firm and its contractual outsourcing partner is an evolutionary one and it may evolve in harmful ways. In such circumstances it may be prudent to maintain a number of suppliers and to rotate the contracts around to reduce dependence on any particular supplier. With each supplier there is a need to monitor and control the outsourced activities to prevent loss of proprietary knowledge that may be vital to the firm's competitive advantage.

**Further alleged pitfalls**

Outsourcing can be performed domestically. It is when outsourcing is shifted offshore that concerns arise. India is the most preferred location for much of the offshoring, because of its cost competitiveness and availability of talented personnel, followed by China and the Philippines. Other popular offshoring locations are Mexico, Poland, Hungary and Vietnam. Direct foreign investment and the establishment of subsidiaries of multinational corporations has long been an avenue to source foreign workers and other resources. The current focus, however, is not on subsidiaries but on outsourcing services directly to foreign firms or teams of specialists that can be assembled to fulfil certain contractual undertakings.

Farrell (2005: 680–681) estimates that every dollar of spending that American firms transfer to India generates $1.46 in new wealth of which 33c remains in India. **American firms save 58c for every dollar of spending on jobs they move to India.** These cost savings can then be used for firm expansion, or for R & D or marketing purposes, to increase staff remuneration, or the benefits of the increased profits can be returned to the shareholders and taxpayers. These are the positives but the headline debates concern the impacts on the jobs of locals that are exported. While US offshoring only accounted for 4 percent of all layoffs, 31% of those whose jobs were displaced by trade were not fully re-employed and 55% of those that were re-employed were at best working for only 85% of their former wages. This then raises the issue of whether enough, or appropriate, trade adjustment assistance and retraining is made available to those who are adversely impacted by offshoring.

A further potential concern is illustrated by Bettis et al (1992:11) relating to the "hollowing out" of American industry and in particular the decline of Western manufacturing. **Key components or complete finished products can be offshored at a delivered cost 10 to 60% below the cost of in-house manufacture**. These authors worry about the surrender of core technological capabilities. Outsourcing decisions cannot be easily reversed. The strategic intentions of the foreign suppliers may be to become potential competitors in the future. Foreign supplier may acquire proprietary design information and if the relationship sours they may use the acquired process and product technology to compete with the Western firm. The suggestion is that outsourcing should be focused on areas far removed from strategic competencies because as outsourcing decisions become closer to core competencies, the strategic risk increases.

An interesting site to explore is: http://www.freelancer.com/ which advertises that one can "Hire freelance programmers, web developers, designers, writers, data entry and more at a fraction of the cost on the World's Largest Outsourcing Marketplace". If the process is carefully managed, rather than home base capabilities erosion it can lead to domestic capabilities upgrading. As more knowledge-intensive and proprietary tasks, including R & D and innovation, are being outsourced local firms can tap into the global talent pool that utilizes skill sets not readily available locally or that are very expensive.

Capabilities are being created through learning from provider's expertise and inter-firm complementarity. To India might be outsourced the labour intensive software coding and to the Philippines the front-end creative design of the User interface. Web designers in the Ukraine or the Philippines can assemble teams of designers very quickly to produce functioning corporate web sites in a week and charge $50 an hour compared to $250 an hour in Sydney, Australia. Australians then take on more of a coordinating/management role, assuring quality outcomes and managing clients. They outsource the technical expertise but not the client/customer relationship.

Another example involves an Australian Skin care company that produces at the top of the product range – a high quality market leading product. The packaging and printing is outsourced to China but the Australian company has to be continually vigilant to protect its intellectual property rights and the quality of the ingredients that are being used. Organic materials cost 7–8 times the cost of synthetics and 2–3 times the cost of non-organic natural ingredients. Foreign firm were found to not be using the contracted ingredients. Clearly the incentive to cheat is enormous when 1kg of regular rose oil was $15 but it cost $15,000 for the organic ingredient.

**Implications**

Offshore outsourcing of services (such as innovation, software development and engineering) can be viewed as welfare-enhancing international division of labour taken to the next stage but it is not without its challenges and opportunities. Firms may need to outsource to upgrade their capabilities and to innovate to survive and grow in the international marketplace. There will be winners and losers. There is a collective responsibility in a humane society to soften the blows of trade restructuring via adjustment assistance. Education systems will be challenged to provide reskilling opportunities but also to focus on creating front office client focused management skills in a context where back-office jobs that can be delivered electronically will be in less demand locally. The challenge will be how to manage the process in an active and not passive way. If it is not managed strategically capabilities and (net) jobs will be lost. Foreign suppliers will expand their capabilities and encroach on core competencies of the buying firm. It might be remembered that Samsung started off as a cheap contract manufacturer. This is particularly the case where foreign firms are agents of the State that heavily subsidize them to gain competitive advantage. Once capabilities are outsourced they are hard to get back – the process can not easily be reversed. If IT programmers in India work at a small fraction of local salaries, then local technical competence may dwindle with fewer local IT graduates as local universities drop these teaching programs.

The key is to actively seek out equity and non-equity based foreign collaborations that provide access to talented personnel and technology – a process of strategic asset seeking. This involves identifying and accessing technologies, competencies and knowledge around the world and using this to create innovative products, services and processes. By relocating specific business process abroad there are cost savings and knowledge acquisitions (Roza 2011). The initiating firm, however, has to call the shots. They need to protect their intellectual property rights, maintain quality control, diversify suppliers and keep their main sources of competitive advantage secure from imitation and replication. The key end-client relationships have to be protected and the overall process management kept in the hands of the initiating firm.

What this case study reveals is that attempts to minimize costs (in the short and long run *and* fixed and variable) are at the forefront of modern corporations in a global context.

**References**

Bettis, R.A., S.P. Bradley, and G. Hamel. 1992. Outsourcing and industrial decline. *Academy of Management Executive*. 6:1 pp. 7–22.

Blinder, A.S. 2006. Offshoring: The Next Industrial Revolution? *Foreign Affairs* 85:2 March/April pp. 113–128.

Contractor, F.K., V. Kumar, S.K. Kundu amd T. Pedersen (2010), Reconceptualizing the Firm in a World of Outsourcing and Offshoring: The Organizational and Geographical Relocation of High-Value Functions, *Journal of Management Studies* 47:8 Dec., pp. 1417–1433.

Farrell, D. 2005. Offshoring: Value creation through economic change. *Journal of Management Studies* 42:3 May pp. 675–683.

Lewin, A. 2012. *Global Sourcing of Business Services: Key Findings and Trends from ORN Research*. The 2012 Outsourcing World Summit, Florida. Feb. Available at http://www.fuqua.duke.edu/offshoring/research/research_findings/

Roza, M., F.A.J. van den Bosch & H.W. Volberda. 2011. Offshoring Strategy: Motives, functions, locations, and governance modes of small, medium-sized and large firms. *International Business* Review. 20. pp. 314–323.

Youngdahl, W. And K. Ramaswamy (2008), Offshoring knowledge and service work: A conceptual model and research agenda, *Journal of Operations Management* 26 pp. 212–221.

# 4 PERFECT COMPETITION

Economists divide the product markets into four categories; perfect competition, monopoly, monopolistic competition and oligopoly. The following four chapters discuss characteristics and behaviour of producers in these markets and at each stage an attempt is made to evaluate markets performance in terms of social benefits and costs.

**Perfect Competition**

A perfectly competitive market is characterized by the following features:

1. Infinite number of small buyers and sellers such that each market participant has no significant effect on the market price.
2. Identical products. In a perfectly competitive market all of the products are perfect substitutes, there is no scope for product promotion or product differentiation.
3. Perfect mobility of the factors of production including labour and capital in the long run.
4. Perfect knowledge of the availability and the quality of products.

The above features are idealistic and do not exist in the real world. The existing markets in the real world may have some features of a perfectly competition. A perfectly competitive market is model that can help to evaluate the degree of competition in real world markets.

**Total, Marginal and Average Revenues**

In a perfectly competitive market a seller can sell all of his produce at the same price. This means an individual produces (seller) faces a horizontal demand curve at the market price. The demand curve facing a competitive producer is also his average revenue (AR) since price is the average revenue. The AR and the marginal revenue (MR) of a perfectly competitive seller are the same because when average is constant, marginal and average are equal. The total revenue (TR), AR and the MR are presented in equations 4.1, 4.2 and 4.3.

$$TR = P \times Q \tag{4.1}$$

Where P and Q are price and quantity respectively.

$$AR = TR / Q = PQ / Q = P \tag{4.2}$$

$$MR = \Delta TR / \Delta Q = P \tag{4.3}$$

The MR is change in TR as a result of a unit sold. MR = P because all of the units produced are sold at the same price. The TR, the demand curve (MR and AR) of a perfectly competitive producer are plotted in Figures 4.1 and 4.2. The total revenue curve is a straight line passing through the origin. The slope of TR is MR which is constant.

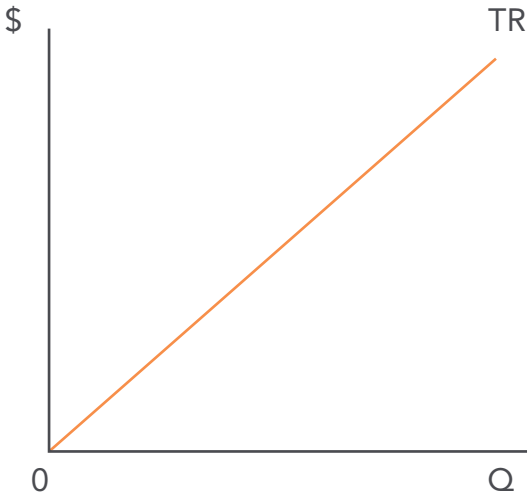


**Figure 4.1** Total Revenue

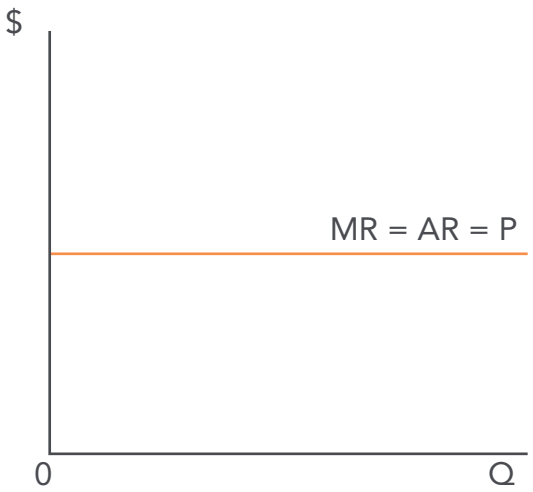


**Figure 4.2** Average, Marginal Revenue

In Figure 4.2 the demand curve facing a competitive producer is infinitely elastic. Because products are identical, sales drop to zero if the price is set slightly above the market price.

### Short Run Equilibrium

The short run equilibrium of the firm is the quantity of output that maximizes the firm's profit. Profit is the difference between total revenue and total cost. Both TR and TC depend on the quantity of output produced.

$$\pi = TR\ (Q) - TC\ (Q) \tag{4.4}$$

In equation 4.4 $\pi$ is the profit and TC includes both explicit and implicit costs of production.

To maximize profit:

$$d\pi / dQ = dTR / dQ - dTC / dQ = 0 \tag{4.5}$$

$$MR = MC \tag{4.6}$$

In equation 4.6 the necessary condition for profit maximization is the equality of price and MC (under perfect competition, price and MR are equal).

The sufficient condition for the profit maximization is:

$$d^2\,\pi / \mathrm{d}Q^2 < 0 \;\text{ implying that } \mathrm{dMR} / \mathrm{d}Q \;\; < \mathrm{dMC} / \mathrm{d}Q \qquad (4.7)$$

Equation 4.7 requires that the slope of MC to be greater than the slope of MR. Since under perfect competition the slope of MR is zero, the slope of MC at the profit maximization output should be positive.

The profit maximisation output is presented in Figure 4.3.

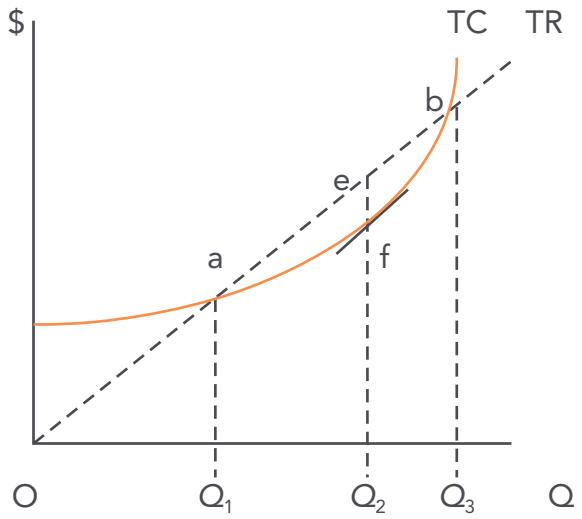


**Figure 4.3a** Profit Maximization

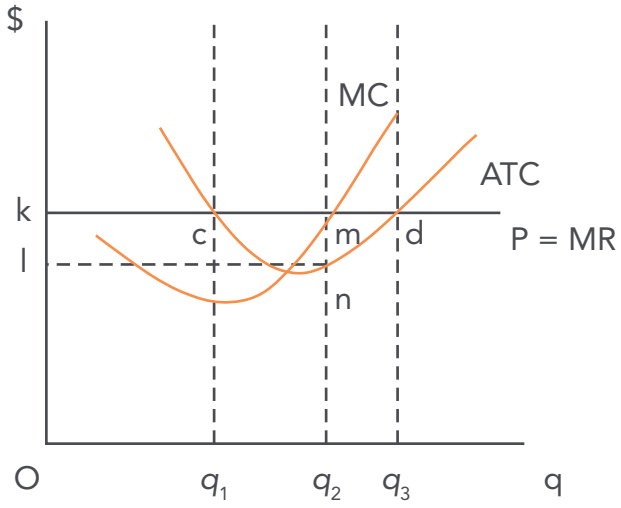


**Figure 4.3b** Profit Maximization

In Figure 4.3a profit is positive between points a and b where TR > TC and also in Figure 4.3b between c and d where P > ATC. In Figure 4.3a profit is maximized at the output $OQ_2$ where slope of TR = slope of TC. This point is projected in Figure 4.3b at the output $oq_2$ where MR = MC and the slope of MC is positive.

In Figure 4.3b at the output $oq_2$, TR is the area of the rectangle okm$q_2$, TC is the area oln$q_2$ and the profit is lkmn. The equivalent of lkmn in Figure 4.3a is the distance ef.

### Short Run Shut-down Point

In the short run it is possible for the firm to make losses, that is when TC > TR. However, in the short run the firm is better-off to continue production if TR > FC. In the short run the FC has to be paid and cannot be changed. The firm should remain in business if it can compensate for part of its Fixed cost. This situation is illustrated in Figure 4.4.

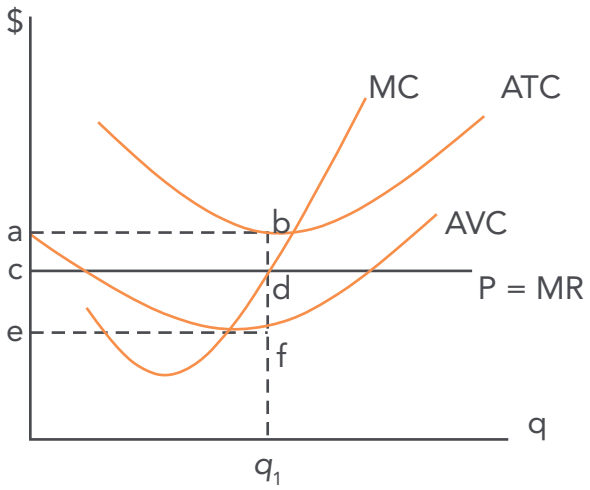


**Figure 4.4** the Shut-down Point

At the profit maximizing output $Oq_1$ the firm is making losses equal to the area cabd because P < ATC (TR < TC). However, at $Oq_1$ the fixed cost is eabf which is payable in the short run even if the firm shuts down. If the firm continues production the loss is less than fixed cost because some of the fixed cost is recovered by the revenue. When TR < TC the firm is better- off to continue production as long as P > AVC or TR > TVC. In the short run when P > AVC, the producer sets the price equal to MC and produces different quantities of output. Accordingly, the portion of MC curve that lies above the minimum of AVC is the short run supply curve because that section of the MC shows various quantities produced at different prices. The industry supply curve of the firms in a perfectly competitive market can be derived by horizontal summation of the supply curves of all of the producers in the market. The equilibrium market price is determined at the intersection of the industry supply and market demand curve. Each producer has no influence on the market price and produces where the market price is equal to the marginal cost of production. The equilibrium market price changes as a result of shifts in market demand or industry supply curves. The market and an individual firm under perfect completion are presented in Figure 4.5.

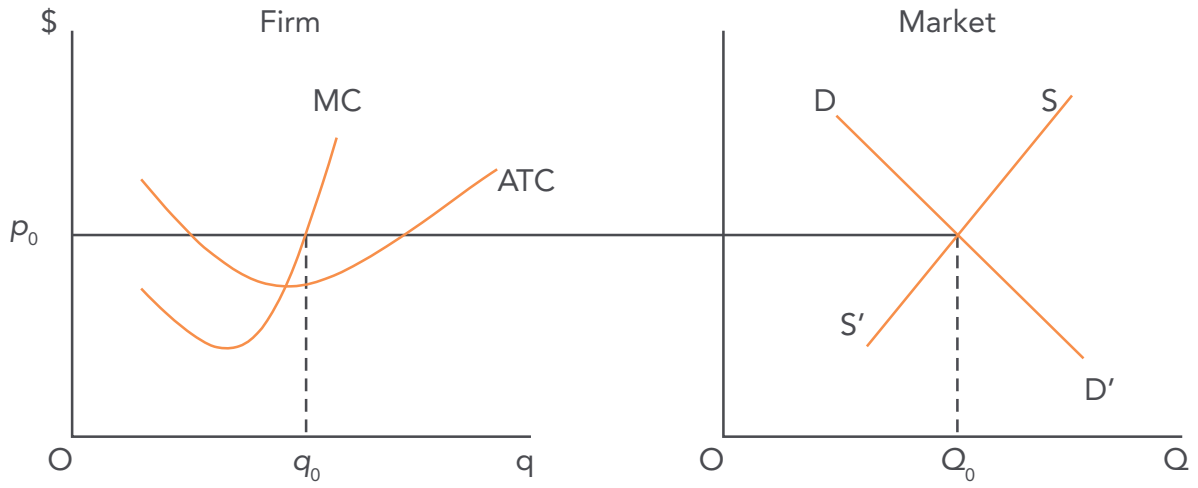


**Figure 4.5** A Representative Firm and a Perfectly Competitive Market

In Figure 4.6 the equilibrium price $Op_0$ is determined in the market at the intersection of DD' and SS'. A representative firm takes this price and produces $Oq_0$ where market price is equal to the firm's MC curve. By producing $Oq_0$ the firm maximizes profit and is in short run equilibrium. At $Oq_0$, market price is greater than ATC and the firm is making more than normal profit. Normal profit is equivalent of implicit cost of production which is the opportunity cost of owner's own resources.

It was discussed earlier that the total cost of production includes explicit and implicit costs.

$$\text{TC} = C_x + C_i$$

Where $C_x$ and $C_i$ are explicit and implicit costs of production respectively. The normal profit is $C_i$.

When Profit is zero, $\text{TR} - \text{TC} = 0$ or $\text{TR} - C_x - C_i = 0$ or $\text{TR} - C_x = C_i$.

The last relationship states that after paying for the explicit costs (labour, raw material, etc.), the residual is the normal profit. In this case price = ATC. Accordingly, as shown in Figure 4.6, profit is greater than normal when price > ATC.

### The Long Run Equilibrium

In the long run firms can enter into or exit from the perfectly competitive industry. Motivation for entry or exit is the profit. When TR > TC the profit for existing firms is greater than normal and other firms outside of the industry are encouraged to enter. The opposite takes place when TR < TC signalling the existing firms to leave. The long run equilibrium in a perfectly competitive industry is established when TR = TC or price is equal to the long run average cost. In this case each firm in the industry is making normal profit.

Suppose initially in Figure 4.6 a representative firm is in short run equilibrium making more than normal profit (price > ATC). Existence of more than normal profit encourages other firms to enter to this industry causing industry supply curve to shift to the right and the market price to decline. As long as Price > ATC other firms enter, the supply curve shifts further and the market price falls. When other firms enter average costs of existing firms may remain unchanged, rise or fall. These three cases are called constant cost, increasing cost or decreasing cost industries. In cases of increasing and decreasing costs, short run and long run average costs also shift up or down respectively. Only constant costs industries are presented here. Adjustments to the long run equilibrium of a representative firm and the market is shown in Figure 4.7.

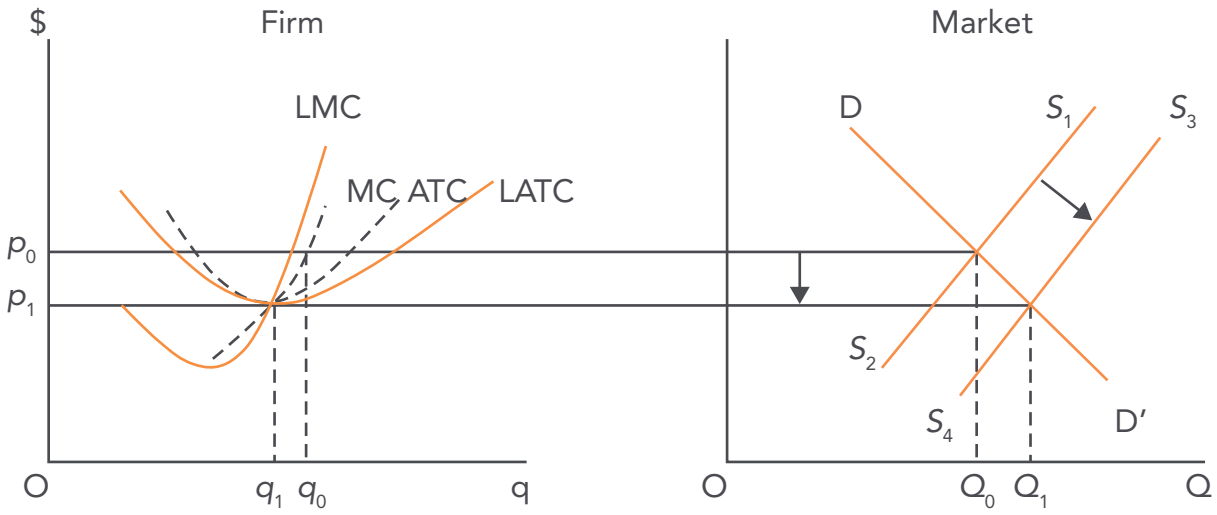


**Figure 4.7** Long Run Equilibrium in a Perfectly Competitive Market

For clarity the short run curves are shown dashed. Initially the representative firm maximizes profit in the short run by producing $Oq_0$ where MC = $Op_0$. The market supply curve shifts to the right as other firms enter the industry. The price falls to $Op_1$ until it is equal to the minimum of LATC. In the long run equilibrium the representative firm is producing $Oq_1$ and the industry is producing. $OQ_1$ In the long run equilibrium Price = MC = ATC = LATC = LMC and each firm in the industry is making normal profit. Long run adjustments take place only when the existing firms in the industry are making more or less than normal profit. The normal profit (the opportunity cost of owner's resources) is the bare minimum that a small competitive firm can earn in the long run.

Examples of markets that have some features and characteristics of perfect competition are agriculture farmers, foreign exchange markets, share markets and mobile phone companies.

### Some Applications of Demand and Supply

Demand and supply analysis is strictly valid only under perfect competition. The supply curve doesn't exist when markets are monopolistic. However, economists and other professional have used the framework without questioning its theoretical or practical validity. The model is the bread and butter of the applied microeconomist. Application of the demand and supply framework will be presented in the following cases: consumer surplus, the effects of sales tax, tariffs and minimum wage.

### The Consumer Surplus (CS)

The CS is related to the demand curve and is defined as the difference between the prices that consumer is willing to pay and the prices that he has to pay. The CS is presented in figure 4.8.

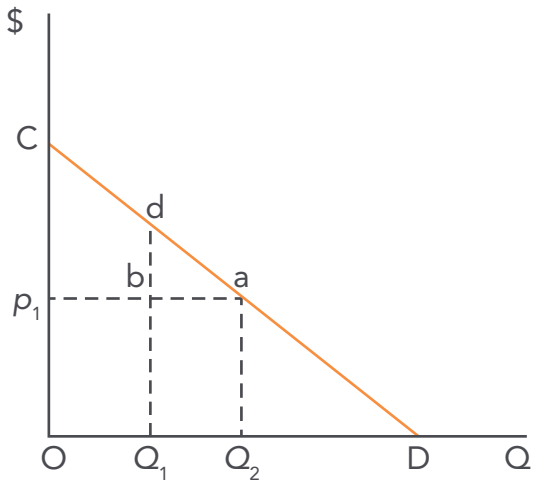


**Figure 4.8** Consumer Surplus

Consider the market demand curve CD and assume that the market price is $Op_1$. At the quantity $OQ_1$, the CS is bd = d $Q_1$- b $Q_1$. The CS at each quantity is always the difference between the vertical distance between the demand curve and the quantity axis and the market price. At $OQ_2$ the CS is the area of the triangle $p_1$ Ca which is sum of the CS of all of the quantities before O $Q_2$.

The CS is related to the welfare of the consumer. Any policy or action leading to a higher CS indicates that the consumer is better-off.

### The Effects of Value added Tax (VAT)

VAT or sales taxes are levied on each unit of output produced. It is like an additional cost of production. Producers attempt to pass on the tax to the consumer in the form of increase in price. Depending on the elasticity of demand, the rise in price may or may not be equal to the full tax – this is called tax incidence analysis. The effects of VAT are presented in Figure 4.9.

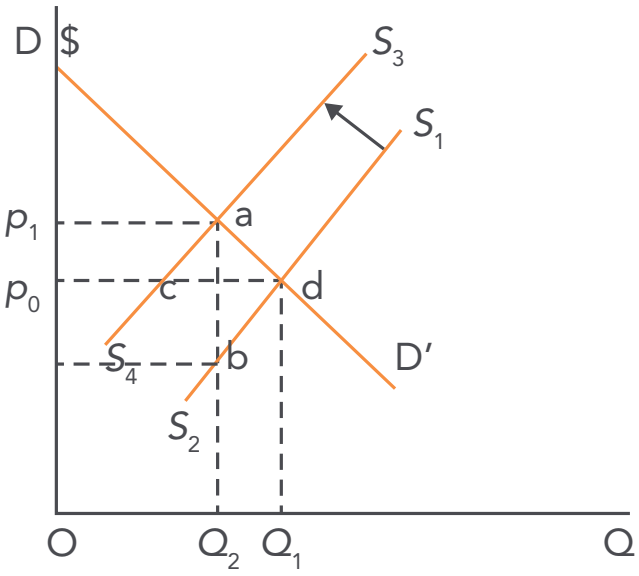


**Figure 4.9** the Effects of VAT

Initially before imposition of VAT the equilibrium is at point d where the market price is $Op_0$. Introduction of VAT causes the supply curve to shift to the left and the market price rises to $Op_1$.

At the new equilibrium, point a, the tax is ab and ac part of the tax is paid by consumer in form of an increase in price. The cb is the producer's share of tax burden. In this case VAT is shared between consumer and the producer. The share of tax burden depends on the elasticity of demand. The lower is the elasticity of demand the larger is the portion of tax paid by the consumer. In case of vertical demand curve (zero elasticity) all of the tax is paid by the consumer.

Imposition of tax leads to a deadweight loss (loss of Consumer Surplus). In the above case the CS before introduction of tax is $p_0$Dd. After tax the CS is $p_1$ Da. $p_0\,p_1$ ad represents the loss of CS as a result of VAT. The dead weight loss in this case is cad because $p_0\,p_1$ ac part of lost CS is transferred to producer's surplus. Most of the government intervention with the market forces causes loss of CS. This topic will be discussed in detail in later chapters.

**Tariffs**

Tariffs are taxes on import. They are designed to discourage imports and promote consumption of domestically produced goods. Assume a small economy that takes the world price as given. At the world price imports are the difference between quantity produced by domestic producers and domestic consumption. Introduction of tariffs increases world's price and reduces consumption and imports.

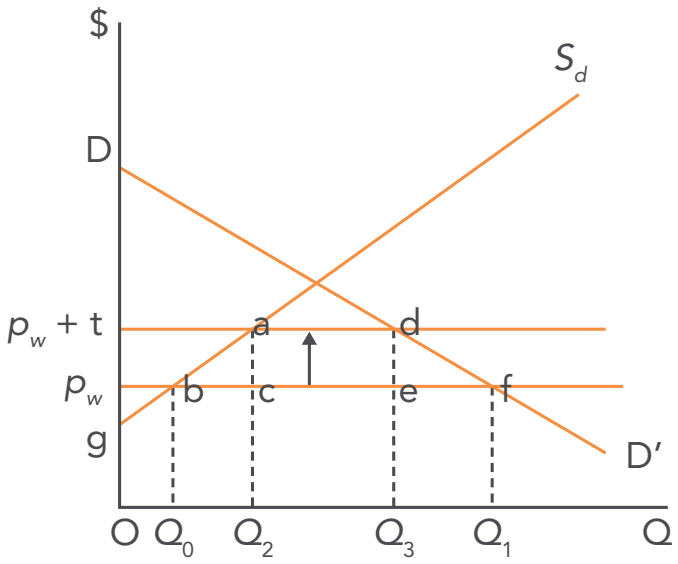


**Figure 4.10** the Effects of Tariffs

In Figure 4.10 DD', $S_d$ and $Op_w$ are domestic consumption, domestic supply and world price respectively. At $Op_w$ the quantity of imports is bf. The introduction of tariffs increases world's price from $Op_w$ to $Op_w$ + t. At the new world price domestic consumption and imports both fall. The government revenue due to tariff is aced. The introduction of tariff reduces CS fromD $p_w f$ to $Ddp_w + t$. The deadweight loss due to the introduction of tariff is abc + def . $agp_w + t$ is the part of CS that is transferred to producer surplus (the difference between price that the producer is willing to receive and the price that is received).

**Minimum Wage**

Often governments or labour unions set the minimum wage in the labour market causing excess supply of labour (unemployment) in the market. The effects of a minimum wage are presented in Figure 4.11.

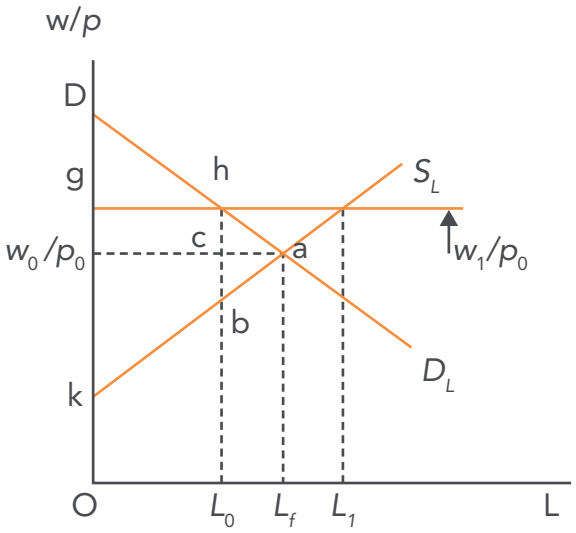


**Figure 4.11** Minimum Wage

In Figure 4.11 $D_L$ and $S_L$ are demand for labour and supply of labour respectively. The minimum wage is set at the real wage $w_1/p_0$ where demand for labour is smaller than the supply of labour. At the minimum wage unemployment is $L_0L_1$. The full employment of the labour market is $OL_f$ at $Ow_0/p_0$.

The introduction of minimum wage also creates deadweight loss equal to the triangle hab. The surplus is transferred from the employers to the labourers. Initially the employers' surplus is $Daw_0/p_0$ and the labourers' surplus is $w_0/p_0$ ak. After imposition of minimum wage employers' surplus is Dhg (the difference between wages that they are willing to pay and the actual wages that they pay) and the labourers' surplus is ghbk. hba is the deadweight loss of minimum wage law.

However, not all labour markets may be competitive. In the case of monopsony a minimum wage may actually not be undesirable. Moreover, the empirical studies of the effects of minimum wages are certainly not clear cut about the overall effects on employment. In some cases a rise in the minimum wage may actually improve employment outcomes.

### Foreign Exchange Market

The global foreign exchange market is one of the most competitive markets. There are very large number of participants around the world and actions of each individual participant has no effect on the market rate. The foreign exchange market is presented in Figure 4.12 where demand and supply for British pound are plotted. The market rate is determined at the intersection of demand and supply curves.

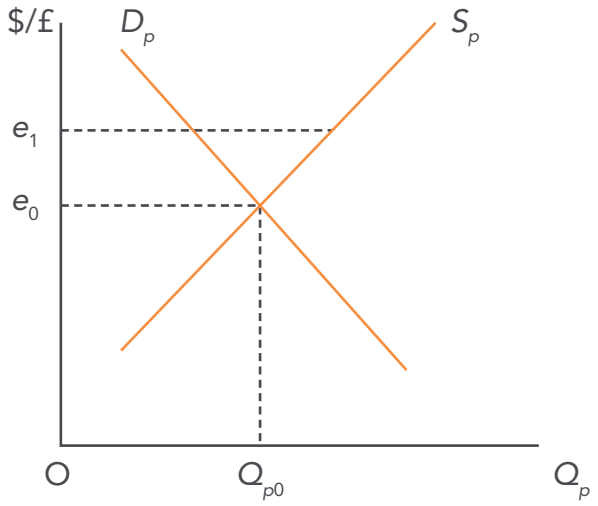


**Figure 4.12** foreign Exchange Market

The foreign exchange rate, $/£, is the price of foreign currency, downward movement means appreciation of pound and upward movement shows depreciation. The demand for pound reflects British exports and capital inflow into Britain. The supply of pounds indicates British imports and capital outflow. The equilibrium exchange rate $e_0$, changes if $D_p$ or $S_p$ shifts.

Sometimes demand and supply for foreign exchange is used to show how a fixed exchange rate system operates. Suppose the authorities wish to fix the exchange rate at $e_1$. In this case the excess supply of pounds has to be purchased by the central bank to prevent the exchange rate from depreciating. The opposite takes place if the central bank attempts to keep the exchange rate fixed below the equilibrium.

### Games that Enhance Learning in Economics

One of the great attractions of a perfectly competitive market is that it achieves economic efficiency. However, efficiency may not be the only consideration – equity and fairness outcomes may also be valued. Is there a tradeoff between these two objectives – equity versus efficiency? Let us look at a game that develops these themes.

The following game – called the *Ultimatum Bargaining Game* – may be useful in stimulating student interest in issues related to equity and fairness in economic decision-making. The aim of the ultimatum bargaining game is to illustrate the concepts of fairness and equity even in 'win-win' situations where both parties can be better off if they agree to do a deal.

In this interactive game, the class is split equally into 'proposers' and 'responders'. So in a class of 30, 15 students take on the role of a proposer and the rest are responders. Proposers are then asked to pair-up with a responder. So there will be 15 pairs. The proposer is then asked to split a hypothetical amount of money (say $10) with the responder in whatever way he or she chooses. For example, the proposer might offer a 70:30 split with the proposer keeping $7 and the responder getting $3. The responder can either accept or reject this proposal. If the offer is rejected then both parties receive nothing. If the offer is accepted then both parties split the $10 as per the offer proposed – $7 for the proposer and $3 for the respondent. Whatever the outcome, in the next round the proposer has to find a different partner to pair-off with. Perhaps play five rounds – so the proposer has five different class members to strike a deal with. Note only one split is offered per trading partner and it is either accepted or rejected – it is a one off deal with no negotiations involved – and then the proposer moves on to the next responder.

Then switch the roles around – proposers become respondents and vice versa – and play another five rounds. After the ten rounds (or whatever number of rounds the teachers thinks feasible in the available time) each student calculates how much money they accumulated – so teachers need to keep track of done deals. Teachers should record all the successful and the *unsuccessful* deals. Then declare which students made the most amount of money, who came second and how much etc.

This is a form of learning through interactive experiments. It is a fun game but it also highlights the extent to which individual behaviour is rational from a strategic point of view, and the role of fairness in bargaining games.

The theory of the rational consumer tells us that the responder will accept any amount above $0, since they would be better off than having nothing. However, in this experiment you will see that narrow self-interest is tempered by notions of fairness and equity. For example, if the proposer offers to split the amount by giving 1c to the responder (and wishing to keep $9.99 for themselves), the responder will most likely see this split as unfair and reject the offer, so that the proposer will end up with $0 as well. This shows how in the ultimatum bargaining game, fairness comes into play and will impact on how the players make their decisions.

We have played this game where 500 students participated but only played two rounds. The average proposal demand was $6.01 for round 1 and $6.09 for round 2. The responder acceptance rate was 64% for round 1 and 62% for round 2. The bar graphs 1 and 2 illustrate the results of round 1 and round 2, showing the number of proposals accepted and rejected at each price. The prices in the experiment are rounded off to the nearest dollar for this analysis.

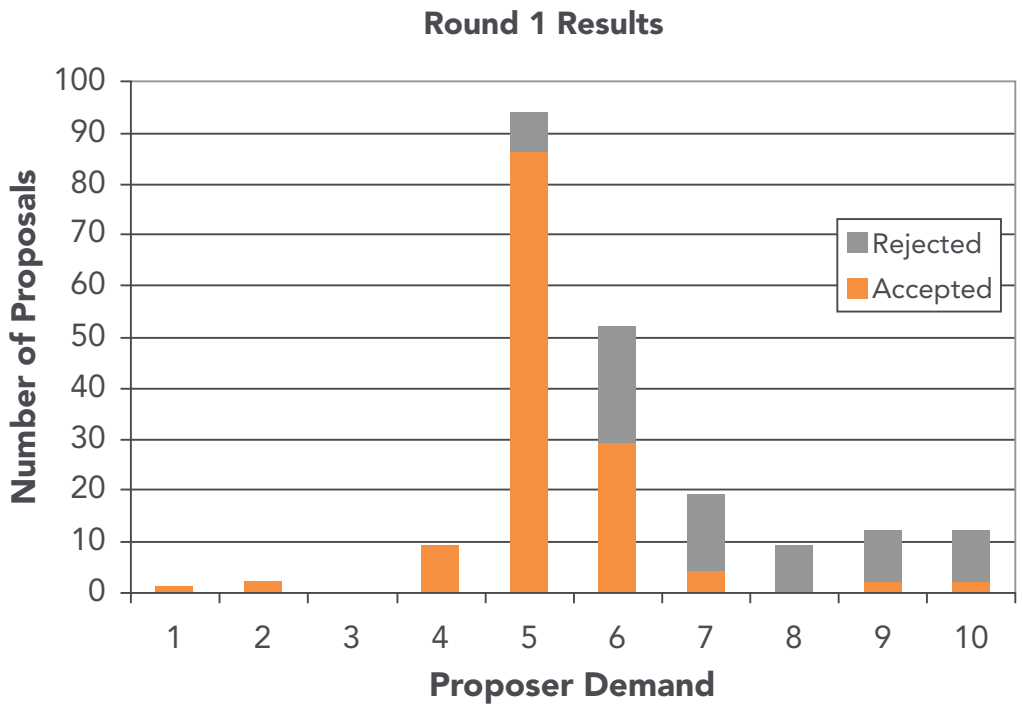


**Figure 1**

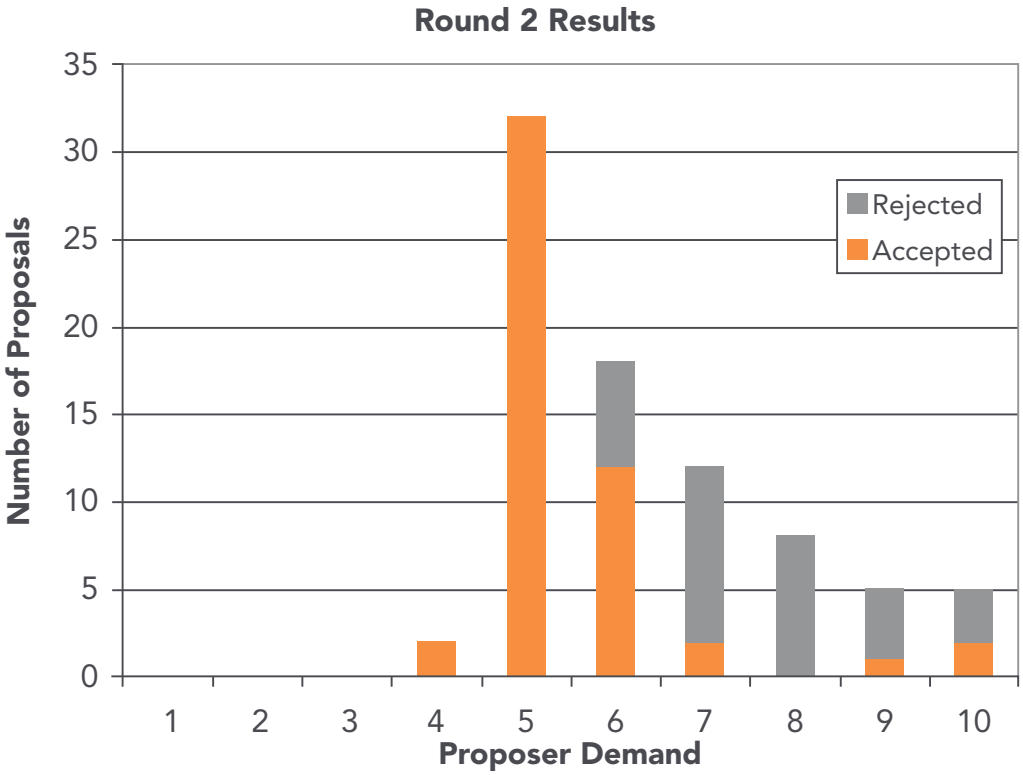


**Figure 2**

The proposals are heavily concentrated to the right of the $5 mark, showing that most proposers wanted more than half of the money, yet $5 was the proposal with the highest frequency, and mostly seen as fair by the responders. All proposals less than $5 – altruistic behaviour? – were accepted. For proposals higher than $5, many were rejected and regarded as "unfair" or "spiteful" and the responders wanted to "punish" this sort of behavior even at a cost to themselves – both parties end up with nothing. In round 2, there are relatively more proposals higher than $5 than in round 1. This can be explained by the large amount of accepted proposals of $5 in round 1, and in response, proposers wanted to try and get more in round 2. This also led to more rejections in round 2, which again shows the consideration of fairness in the responders' decision.

The following student comments show the student-teacher interaction process. The teacher's comments are in italics.

**Student 1**: Interesting little game. I was a proposer and I offered 6-4 twice. It was accepted the first time and rejected the second time. Noting that unless the offer was 10-0, where the responder should be indifferent whether to accept or reject, all other amounts, if accepted, will be of benefit to both parties. I guess it demonstrates how people react to real or perceived injustices. A person would rather, out of spite, have both parties get nothing, than accept an unfair distribution. I think at 5-5, it should be 100% acceptance rate as 1) acceptance will offer both parties benefits and 2) it's a fair distribution. So as the distribution gets more and more unfair, the responder would have to decide whether to accept the unfair distribution or punish the proposer.

I think there must be a tipping point, where at, anything less fair, the responder will reject, and anything fairer, the responder will accept. The trick for the proposer is to find that point and propose that amount from that point on. Too bad the game is only repeated once!

**Student 2**: One thing I just thought of was, does that fact that we are using virtual money affect it? i.e. because we will not get any 'real' world benefits, it's easier for us to reject? Maybe you should try this with real money next time.

*Great comments. You have got out of it exactly what I wanted you to get out of it. Rationally you should accept any offer greater than zero as it is a win-win situation. But justice, equity and fairness issues intrude leading to an equity-efficiency trade-off. That is why equity issues are so important.*

**Student 3**: I guess I am getting a free meal by clicking accept. But I don't like someone else getting more than me. But the game is strongly biased (proposer decides the split) where's the bargaining part of this game? It is just one way bargaining. I think we should be allowed to propose a counter offer for this game to be a proper bargaining game.

*That is called envy. It also means that your utility or happiness depends on your relative position among others. This may well be a characteristic of many consumers but it is different that the simple self-interested individual behavior that we commonly assume. On your second point, I agree.*

**Student 4:** I think the game also depends on the number of times or rounds which you get to offer and accept. For example if there was only one round and the offerer gives me an offer of 9-1 the only rational thing for me to do would be to accept it as I would get nothing in total otherwise and 1 is obviously > 0. Now if there are 2 rounds the whole equation changes. If the guy offers me 9-1 I would decline and both of us will get nothing. If I accepted the offer I would get $1 for the round but it would send my offerer the wrong message. It would make the offerer believe that 9-1 is fair and that I am actually happy with 9-1 and hence he will continue to make another 9-1 offer. Now this is not true because I can actually get more money than just a total of $2, although there is an element of risk. Nothing ventured, nothing gained.

After I decline the first offer the offerer will realise that offering me 9-1 would most likely result in another rejection. Also if he makes an offer of 8-2 he would feel that he would be taking a risk and he actually has much more to lose than me (responder). So there is a high chance that he will hedge his bets and offer 6-4 or 7-3 which I would happily accept. As the number of rounds increases the offerer will have to offer more and more to keep the responder happy. As the number of rounds reached infinity or a very large number the offerings, it should actually become 5-5.

*I agree. This is the difference between one-off games and repeated games.*

**Student 5**: If the sum of money used in the experiment was $10,000 the results would be very different (assuming real money was used). If we were offered 10% of the amount, I think most people in their right mind would take the $1000, no questions asked. I would assume no-one would reject $1000 because of the inequality. So I guess the experiment only applies in consideration of these factors; a computed sum was used, and the sum was only $10.

*Not sure about that. Witness the breakdown of World Trade Organization talks. The rich countries were offering cuts in agricultural protection levels and subsidies. But the developing countries said no – the process was inequitable. The talks collapsed. Poor countries walked away from deals that would have considerably increased their export income.*

What this game indicates is that if the task is presented in an interesting way students will voluntarily join-in and deepen their understanding of concepts and be stimulated to ask searching questions of the material.

## Conclusion

The application of models based on perfect competition can generate insight into many issues and problems. However, one needs to be careful that the analysis is too simplistic and omits important variables (equity, fairness and justice for example) or it is applied to markets where there is market power that convers consumer surplus to producer surplus. It is these non-competitive markets that we now turn to.

# 5 MONOPOLY

Monopoly is a market situation where there is only one single seller, no substitute exists and entry to the market is blocked. A monopolistic market is opposite of perfect competition. Similar to perfect competition, pure monopoly is exceedingly difficult to find in the real world. Knowing characteristic of this market allows researchers to evaluate the degree of monopoly power in the real market situations.

### Average Revenue and Marginal Revenue

A pure monopolist faces the market demand curve. The average revenue curve of a monopolist is downward sloping because the market demand curve is downward sloping.[5] Since AR is falling, in this case, the MR curve is always below the AR. AR and MR curves facing a monopolist are plotted in Figure 5.1.

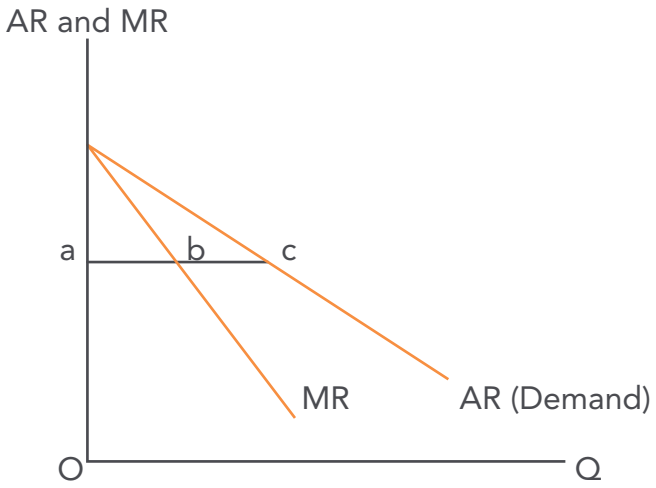


**Figure 5.1** Average and Marginal Revenues

The MR curve cuts the mid- point of any horizontal distance between the AR and vertical axis. This means that the slope of MR is twice the slope of AR. Consider the following relationships:

$$P = AR = F(Q)$$

$$TR = F(Q) \times Q$$

$$MR = dTR/dq = F'(Q)Q + F(Q)$$

The slope of MR is $F''(Q)Q + F'(Q) + F'(Q)$

The slope of MR is 2 $F'(Q)$ because for a linear MR $F''(Q)$ is zero.

The slope of AR is $d\,F(Q)/dQ = F'(Q)$.

**AR, MR and Elasticity**

It was shown earlier that $MR = dTR/dq = F'(Q)Q + P$ (1)

Dividing both sides of equation 1 by P yields:

$$MR/P = F'(Q)Q / P + 1 \quad (2)$$

The elasticity of demand is $e = dQ/dp \times P/Q = 1/ F'(Q) \times P/Q$

Accordingly 2 becomes: $MR = P(1/e + 1)$ or $MR = P/e + P$

Since $e < 0$, $MR < P$. In case of perfect competition $e = \infty$ and $P = MR$.

The relationship between P, MR and e is useful for comparison of perfect completion and monopoly and also for measurement of the degree of monopoly power. It is customary to write the relationship in form of MR = P – P/e where e is the absolute value of the elasticity of demand.

### Short Run Profit Maximization

A monopolist maximizes profit where MR = MC and the slope of MC is positive. The profit maximization output of a monopolist is plotted in Figures 5.1a and 5.1b.

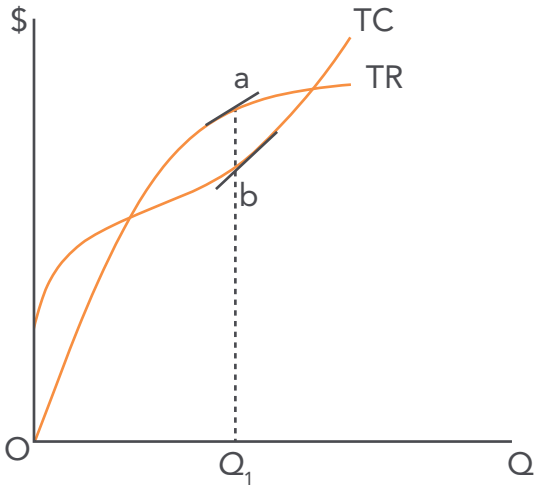


**Figure 5.1a** Profit Maximization

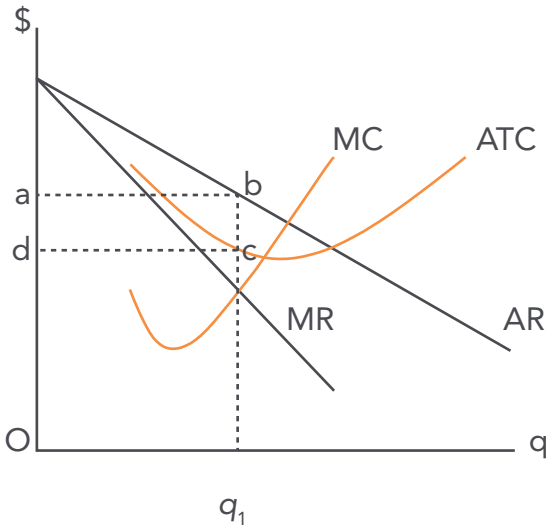


**Figure 5.1b** Profit Maximization

In Figure 5.1a profit is maximized where tangents to TC and TR are parallel, at the output $OQ_1$. The profit is vertical distance ad. In Figure 5.1b $Oq_1$ where MC = MR corresponds $OQ_1$ to in Figure 5.1a. At the profit maximizing output, the price per unit is Oa and the profit is the area of the rectangle abcd. The total cost of production is the area dc$Oq_1$. In this case the monopolist is making more than normal profit. The price Oa and the quantity $Oq_1$ are the only pair of price and quantity that maximizes monopolist's profit, there is no other pair. For this reason the supply curve doesn't exist under imperfect competition. Point b is the only point on the supply curve; it is not possible to trace additional profit maximizing points on the curve.

### Long Run Equilibrium

In the long run the monopolist maximizes profit where LMC = MR. Since the entry into the monopolist's market is blocked, unlike perfect competition, economic profit will not disappear and monopolist is able to earn more than opportunity cost of owner's own resources. The long run equilibrium of a monopolist is presented in Figure 5.2

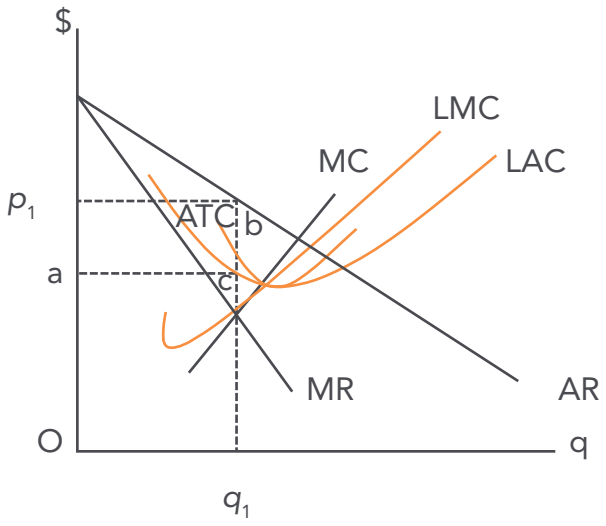


**Figure 5.2** Long Run Equilibrium

The long run profit maximizing output $Oq_1$ is where LMC = MC =MR. At $Oq_1$ the price is $Op_1$ and the economic profit is $ap_1$bc. The economic profit will be maintained because other firms cannot enter into a pure monopolistic market. For this reason sometimes it is argued that the existence of monopolistic enterprises contributes to the inequality of income distribution because firms earn and maintain large profits.

**Price Discrimination**

A monopolist can sell its products in separate markets at different prices. For example selling the same product at home and overseas at different prices or electricity and gas companies charge different rates for businesses and individual households. For effective price discrimination, markets have to be separated and the elasticity of demand in different markets has to be different. The separation of markets is obvious otherwise consumers buy in the cheap market and sell in the more expensive market. The reason for different elasticity will be discussed later.

The monopolist must distribute output in separate markets such that marginal revenues in different markets are equal. This allocation of products among markets maximizes total revenue of the monopolist.

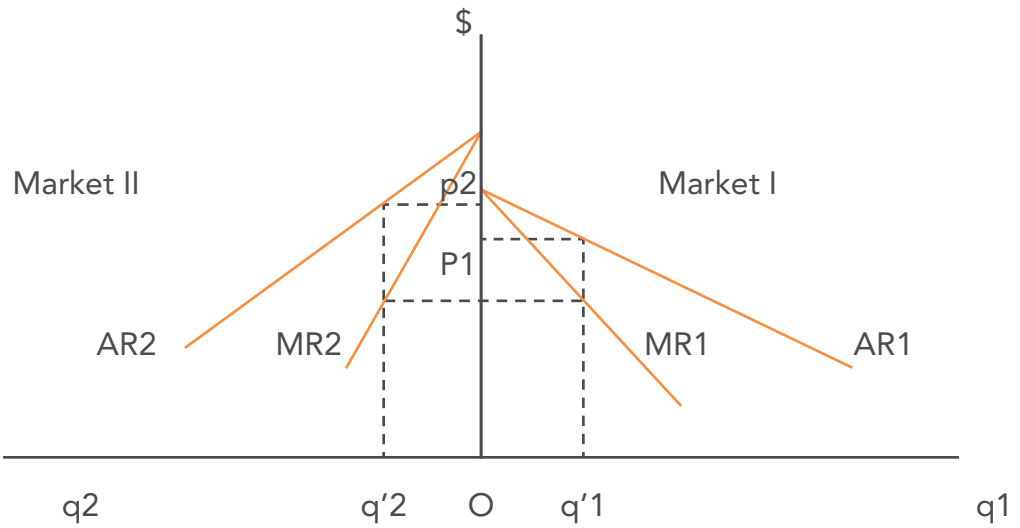


**Figure 5.3** Price Discrimination

In Figure 5.3 there are two markets where AR1 and AR2 are drawn such that elasticity of demand at any price is different in two markets. By equating marginal revenue in both markets the monopolist is charging higher price in market II than in market I.

Suppose the monopolist produces Oq'1+ Oq'2 and wishes to distribute the total product in two markets. In order to maximize total revenue he must sell Oq'1 in market I and Oq'2 in market II such that marginal revenues in both markets are equal. If he disturbs this distribution his total revenue will decline. Suppose he transfers one unit from market I to market II. The loss of revenue is MR1 whereas the gain in revenue is MR2. MR2 > MR1after one unit is transferred from Market I to market II.

For a successful price discrimination elasticity of demand has to be different in different markets.

Earlier the relationship MR = P – P/e was developed. Since MR1 = MR2;

p1(1 – 1/e1) = p2(1 – 1/e2). If p1 is different from p2, for the equality to hold e1 must be different from e2. In the above case p1 < p2, e2 < e1. This means that the monopolist can charge a higher price in the market where the elasticity of demand is low. Generally, in markets with low elasticity fewer substitutes are available and the monopolist can charge higher prices. Those firms that sell domestically and also export often charge lower prices overseas where competition is more severe and larger number of substitutes is available.

**Price Discrimination: A case study**

Here we present a nice question that brings in innovative pricing strategies, the idea that there is consumer surplus available to be captured, calculations of TR and TC, and price discrimination.

*Information*: Warner Bros. owns *The Fugitive* and *Free Willy*. After the release of the two movies, the Company is ready to sell the two DVDs to the mass market. Warner Bros. unit costs for the DVDs are $5 and the stores that sell the DVDs charge a markup of $9.

Assume the market is as follows:

100 people would pay $20 for *The Fugitive* but had no interest in *Free Willy*

100 people had just the opposite preference

100 people said they would buy both movies at $20 each

100 people said they liked both movies but that $20 was too expensive, but would buy both at a lower price of $17.50

a. Warner Bros. is considering two possible prices – $20 or $17.50 – which price should it charge? Why?

Answer:

If it charges $20 then Q=400 so TR = 8000. ATC = 14 so TC = 14 × 400 = 5600. So Profits = TR – TC = 2400

Another way to do it is to say that profit per copy = 20 – 5 – 9 = 6 so 400 × 6 = 2400.

If the P = 17.50 then profit per copy is 17.50 – 5 – 9 = 3.50. Q = 600. So profit is 600 × 3.50 = $2100.

So make $300 more if charge $20.

b. Would the company make more money if it charged $20 for each DVD but gave purchasers of either DVD a coupon for $5 off the purchase of the other DVD?

Answer:

Sales would now be 200 at price \$20 AND 200 at price \$35. At price \$20 the profit per copy is \$6. At price \$35 (for both) the profit per copy is \$35 – \$5 × 2 – \$9 × 2 = \$7. So total profits are \$6 × 200 + \$7 × 200 = 1200 + 1400 = 2600.

So it would earn \$200 more than if it just charged \$20.

c. Why does the studio come out ahead if it cuts the price of both DVDs together by \$5 but losses if it cuts the price of each DVD separately by \$2.50?

Answer:

Because a package discount gives fewer people the discount. The package price discount attracts a new group of customers to the market without having to extend the discount to all existing customers (only some).

d. Would the discount package work if the two DVDs were *Star Wars* and *The Empire Strikes Back*?

Answer:

No. You would either buy both or none at all. Those that bought would just get a general price cut so everybody would get the discount. You would not be able to restrict the discount to certain groups because of different preferences.

e. What problem would the company face if people starting selling the coupons i.e. a market for coupons emerged?

Answer:

Now things start getting interesting. If a market emerged the first 2 market segments could now start selling the coupons to customers that wanted to buy the other DVD. Say that a purchaser of The Fugitive sells the coupon to Free Willy (which it does not want) for $3 (with a face value discount of $5). Then the Free Willy customer only has to pay effectively $20 – discount $5 + $3 = $18. So the first customer paid 20 – 3 = 17 and the second customer paid 18. Before a market emerged they both paid $20. So Warner Bros. loses revenue. That is why these schemes have all sorts of restrictions to prevent a market emerging – non transferable, identity checks etc.

**Monopoly and Perfect Competition Compared**

It is often argued that a perfectly competitive firm produces larger quantity at a lower price and is more efficient than a monopolist. Accordingly, the society is better-off with more competitive firms. There are pro and cons of these arguments. Those that support perfect competition based their argument on the long rum equilibrium of the two market situations. Consider Figure 5.4.

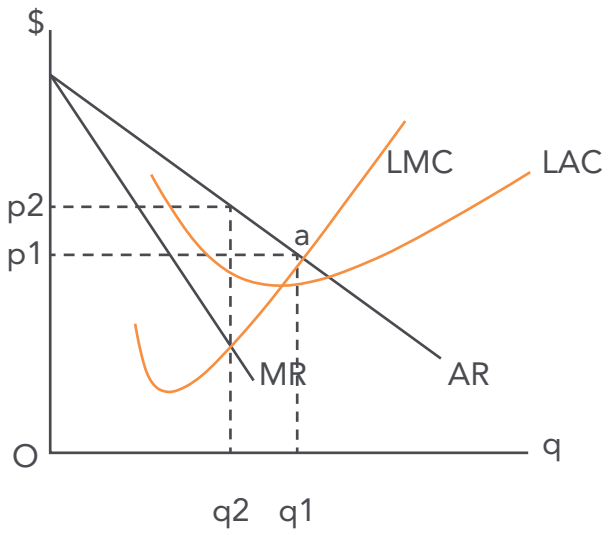


**Figure 5.4** Monopolies and Competition

Assume cost functions of a competitive firm and a monopolist are the same. The monopolist produces Oq2 and charges Op2 where MR = LMC < p2. The competitive firm produces Oq1 at point a where LMC = AR and charges Op1. A competitive firm produces more and charges less. Consumers are better-off under competitive condition. Furthermore, a competitive firm is more efficient because in the long run produces at the minimum of LAC, a monopolist average cost of production is necessarily higher. A competitive firm is more equitable because in the long run earns only normal profit, a monopolist maintains large profits in the long run.

Those in support of monopoly argue that monopolists are large enterprises with access to vast quantity of financial resources. These firms conduct research and development and improve the quality of product for the benefit of consumers that a small competitive firm is incapable of doing so. A small competitive firm, particularly in agriculture, cannot survive in the long run without government subsidies. However, a monopoly firm with large financial resources out of undistributed profit is self sufficient and can remain in production for a long time.

The deadweight loss associated with monopoly is presented in Figure 5.4.

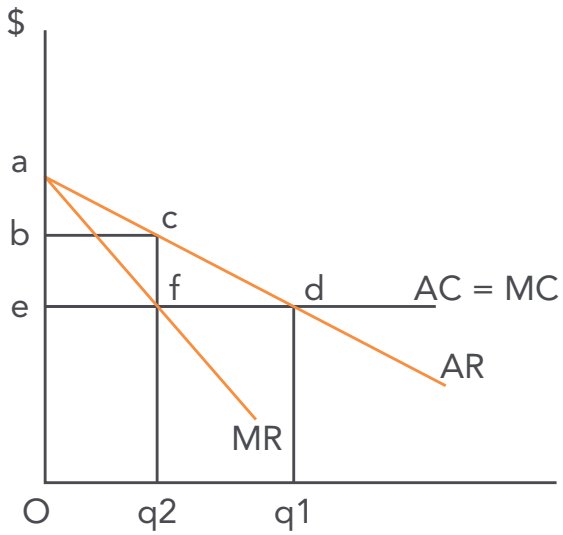


**Figure 5.4** Monopoly Deadweight Loss

In Figure 5.4 for simplicity it is assumed a constant average cost where average cost and marginal cost are equal. The monopolist maximizes profit by producing Oq2 and charging cq2. The consumer surplus is abc and the producer surplus is ebcf. A perfectly competitive firm produces at point d where price is equal to MC. In this case the consumer surplus is ade. The deadweight loss associated with the monopoly is the triangle fcd.

## Sources of Monopoly

1. Control of essential inputs such that no other firm can produce a particular service or commodity.
2. Government regulations licensing a firm to be a single producer of a commodity or service.
3. Patents allowing a sole producer to produce certain goods for a limited period of time.
4. Natural monopoly where cost structure of a firm allows them to set the price so low that it is not worthwhile for other firms to start production of that commodity.
5. Climate and geographical location allowing a commodity being produced in a particular region; French Champagne and Scotch whisky.

## Natural Monopoly

It was argued earlier that natural monopoly is a market situation where one firm experiences falling average cost for extended quantities of output. This firm is able to set the price so low that it is not worthwhile for the other firms to enter into this industry. Utility companies such as electricity, telephone and gas companies are examples of natural monopoly. Price setting by a natural monopolist is presented in figure 5.5.

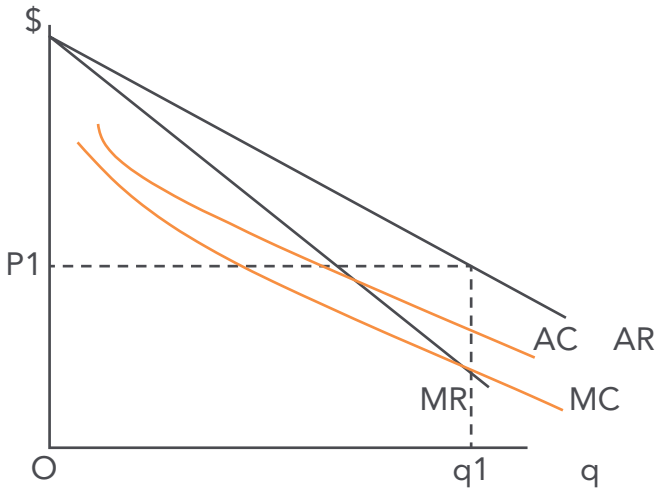


**Figure 5.5** Natural Monopoly

In Figure 5.5 it is assumed that average cost is failing for large quantities of output and MC remains below average cost. The monopolist produces Oq1 and charges Op1 per unit. This price is too low for other produces to enter and the monopolist serves the entire market. In utility companies (electricity, gas, water) usually capital investment is high. Once heavy equipment are installed, it is possible to serve a large community with no extra cost and as the quantity rises average cost falls. In these cases, because of high initial capital investment and low price, new firms are discouraged to enter.

**Regulation of Monopoly**

Monopoly high profit and high price can be regulated by authorities imposing price control. If price is set equal to average cost, monopolist's super normal profit disappears and consumers benefit from lower prices and increased surplus. This type of regulation reduces monopolist's financial resources and prevents engagement in research and development for improvement of quality. The next best solution is to set the price equal to marginal cost which is close to competitive pricing. In several countries governments impose price control on natural monopolist that produce essential services such as electricity, gas and water. The price control is usually in form of marginal cost pricing.

**A case study: Monopoly and industrial innovation**

There does not seem to be widespread agreement among economists about the appropriate theoretical treatment and analysis of industrial innovation. This is particularly disturbing as technical progress is an important, perhaps major, component of the growth of productivity in economies.

The link between market structure and industrial innovation is an interesting one. It is important to examine the interaction between the number and size of incumbent firms in an industry (and any potential entrants waiting in the wings) and technology performance, in terms of process and product innovation. Here are some questions to ponder:

What sort of market structure is most likely to generate a rapid rate of innovation?

Does 'too much' competition stifle the incentive to innovate by making it more difficult to appropriate the returns to research?

Do firms have to be of sufficient size to engage in a significant innovative effort?

Should governments promote mergers of firms or takeovers in particular industries to improve their long-term competitive position in the international market place?

Is it true then that the market structure that attains static efficiency may not typically be well suited to achieve dynamic efficiency?

There are now several excellent surveys examining the relationship between market structure and industrial innovation. A number of these studies have focused on the so-called Schumpeterian hypothesis that large firm size, diversification or monopoly power (in terms of high market concentration) is conducive to innovation and technical progress. A wide range of empirical interpretations of this hypothesis have been formulated. Studies have used research and development (hereafter R&D) spending, R&D intensity (R&D/sales) and the number of personnel engaged in formal R&D activities as proxies for inventive activity. Leading output measures are patents awarded, 'important' patents, and sales of new products. None of these proxies are completely satisfactory. The following results seem to emerge:

1. The smallest manufacturers do little formal R&D while medium size manufacturers support a volume of R&D proportional to their sales. Nevertheless, small firms contribute disproportionately to the creation of new products and processes, although larger firms often develop these inventions for commercial utilisation. The largest firms do not carry out a large share of the relatively risky R&D or of the R&D aimed at entirely new products or processes.
2. Within industries, R&D employment or spending rises proportionately with firm size up to a threshold. A similar pattern emerges with innovative outputs as measured by patents awarded.
3. The evidence on diversification is ambiguous. Beyond a very narrow focus, diversification into related products is favourable but very broad diversification is of little benefit.
4. Also subject to a threshold, seller concentration has a positive but weak impact on research activity. A firm whose monopoly profits are secure is least likely to innovate, takes the longest time to develop an innovation and is most likely to discontinue development if preceded by a rival. New entrants or firms earning only normal profits are the most vigorous innovators. However, the more deeply a rival innovation cuts into a firm's current profits, the greater the incentive to innovate as a defensive measure.

The evidence shows the search for an optimal firm size for invention and innovation to be misguided. There is room for firms of all sizes. While a little bit of bigness is good for invention or innovation, further size adds little or nothing. There is also no simple relationship between market structure and technological progressiveness. High concentration levels tend to discourage innovative output and hence barriers to new entry should be kept at modest levels so that established industry members can be continually exposed to the threat of entry by technically audacious newcomers.

Finally, other variables such as technological opportunity, degree of product differentiation, and the extent of industry rivalry are likely to condition and interact with market structure in influencing industrial innovation.

# 6 MONOPOLISTIC COMPETITION AND OLIGOPOLY

## Monopolistic Competition

Monopolistic competition is a market situation where there is large number of small buyers and sellers, products are slightly differentiated and the entry to the market in the long run is open. Retail sales such as food, clothing and footwear all are examples of monopolistic competition.

Product differentiation is an important feature of this market. Perhaps, packaging, brand name or other superficial differences exist between the products. Sellers attempt to distinguish their products by advertising.

## Short Run Equilibrium

Since products are close substitute the demand curve (average revenue) facing each seller at each price is highly elastic, the curve is relatively flat. Profit maximization output of an individual seller under monopolistic competition is presented in Figure 6.1.

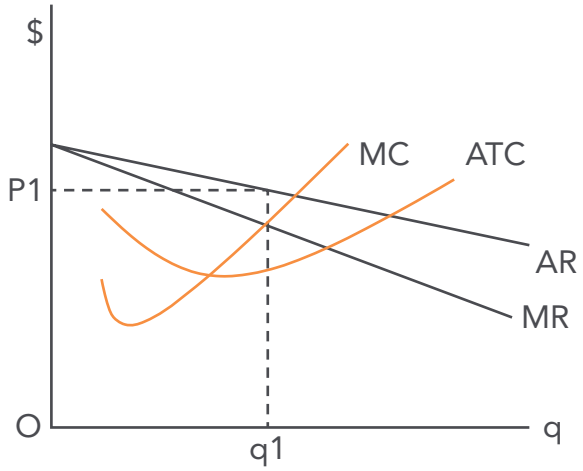


**Figure 6.1** Short Run Equilibrium

The firm in Figure 6.1 maximizes profit by producing Oq1 and charging Op1. By producing Oq1 the firm is earning more than normal profit. Since entry into the market is free, the existence of super normal profit encourages other firms to enter the market and produce similar but differentiated product.

### Long Run Equilibrium

As other firms enter the market, the existing firms experience a smaller market for their product. As a result the demand curve (average revenue curve) of an existing firm shifts down. As long as AR is above the ATC, super normal profit encourages other firms to enter the market causing the demand curve for existing firms to shift down. The long run equilibrium position of a representative firm earning normal is presented in Figure 6.2.

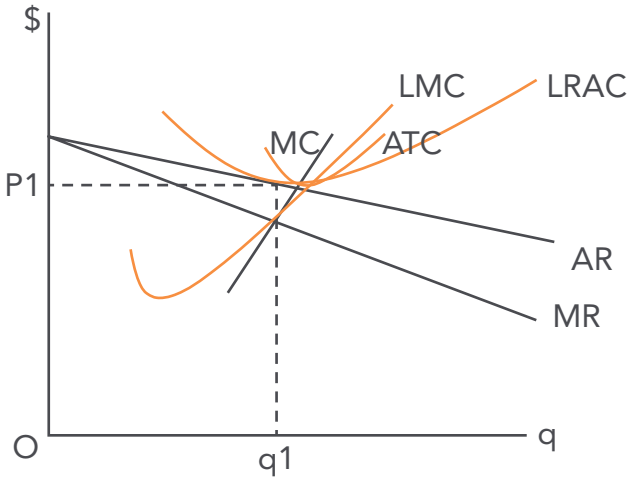


**Figure 6.2** Long Run Equilibrium

By producing Oq1 and charging Op1the representative firm is in the long run equilibrium making normal profit where price is equal to the average cost. The long run position of the representative firm is similar to a firm in a perfectly competitive market except that the equilibrium is not at the minimum of the average cost curve. For this reason it is argued that a monopolistically competitive firm operates with excess capacity. Firms operating with excess capacity, and producing with higher than minimum average cost, remain in the market. Consumers pay prices that are higher that what they could be charged if the firm was producing at minimum average cost. However, it can be argued that consumers benefit from product diversity because of large number of firms in the market producing differentiated products.

**Advertising**

In a monopolistically competitive market firms attempt to distinguish their products through advertising and other forms of product differentiation. Some argue that by informing the consumers, advertising creates new demand leading to higher consumption and employment. Those who oppose advertising argue that advertising is waste of resources because consumers' wants are created by advertising and are not genuine.

Whatever pro and cons, substantial quantity of resources are allocated to advertising in free enterprise systems. When new products enter the market it is essential to introduce them to the consumers. The consumers are free to choose among alternatives but advertising a new product gives consumers a wider choice and not necessarily inefficiency in resource allocation. Since products are close substitutes, by choosing the advertised product, consumers' wants are diverted from the substitutes and hence there is no change in aggregate consumption and overall resource allocation. However, remember our earlier discussion about conspicuous consumption and artificially created 'wants'.

**Oligopoly**

Oligopoly is a market situation where there are a few producers producing homogeneous or differentiated goods. The entry to the market is generally blocked, because of heavy capital investment, collusion of existing firms or government regulations. The auto industry, aeroplane manufacturing and computer operating system companies are example of oligopoly markets.

An important feature of this market is interdependency between firms. Price and quantity of each firm depends on price and quantity of rivals. It is not possible to plot the demand curve facing an individual seller without making assumptions about the re-action of the other firms when one firm sets its price and quantity. There are several theoretical models regarding derivation of the demand curve of an individual producer in an oligopoly market structure.

**Cournot Model**

The earliest solution to the oligopoly situation was offered by the French mathematician Augustin Cournot. The model assumes that there are two firms I and II (duopoly) each producing the same product with identical cost functions. Both firms are aware of the demand curve for its product. Each firm attempts to maximize profit assuming that the output of the other firm remains constant.

Suppose the market demand curve is $P = f(q_1 + q_2)$, where P is the market price, $q_1$ and $q_2$ are quantities produced by firm I and II respectively. $MC_1$ and $MC_2$ are equal because both firms have identical cost functions. Profit maximizing equations for each firm are presented in equations 6.1 and 6.2.

$$\pi_1 = q_1 f(q_1 + q_2) - C_1 \qquad 6.1$$

$$\pi_2 = q_2 f(q_1 + q_2) - C_1 \qquad 6.2$$

Differentiating 6.1 and 6.2 with respect to $q_1$ and $q_2$ and equating each equation to zero and solving $q_1$ for and $q_2$ yields:

$$q_1 / q_2 = df / dq_2 / df / dq_1 \qquad 6.3$$

The right hand of 6.3 is equal to unity because of downward sloping demand curve facing each firm. Accordingly $q_1 = q_2$. The market price can be determined by inserting values of $q_1$ and $q_2$ in equations 6.1 and 6.2.

Producing equal amounts should not be surprising when firms have identical cost functions and facing the same market demand curve. The Cournot solution is a Nash Equilibrium because each firm maximizes profit given what the other firm is doing.

**Kinked Demand curve (KDC)**

The KDC attempts to explain why prices in oligopoly markets are stable. Introduced by Paul Sweezy (1953), the KDC assumes that when one firm reduces price other firms also reduce their price to maintain their market share. However, when one firm increases price others will not react because they can attract the first firm's customers. Under this assumption the elasticity of demand is higher for price increases and lower for price reductions.

The kinked demand curve is presented in Figure 6.3.

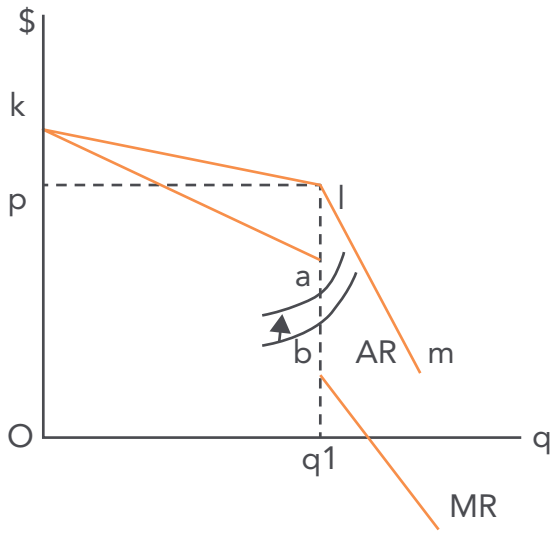


**Figure 6.3** Kinked Demand Curve

The kinked demand curve is klm where the kink occurs at the market price OP. For prices above Op price rises are not followed by rivals, the quantity falls by a large amount making the elasticity of kl segment to be high. Price reductions below Op are followed by other firms leaving a negligible effect on the quantity, low elasticity of demand. The corresponding MR curve is plotted below the AR with a discontinuous segment ab. If the MC curve falls in the ab segment, the profit maximizing price and quantity are Op and Oq1 respectively. In the segment ab increases in cost (upward shift of the MC curve) leave the market price unchanged.

The KDC has been criticized by many economists. The model offers no explanation for determination of market price Op. Furthermore, if MC curves fall in any other segments of MR curve, a rise in cost leads to an increase in price.

### Collusion

Chamberlin (1933) showed that the best solution for oligopoly firms is to collude and act like a monopoly. Chamberlin's solution is presented in Figure 6.4.

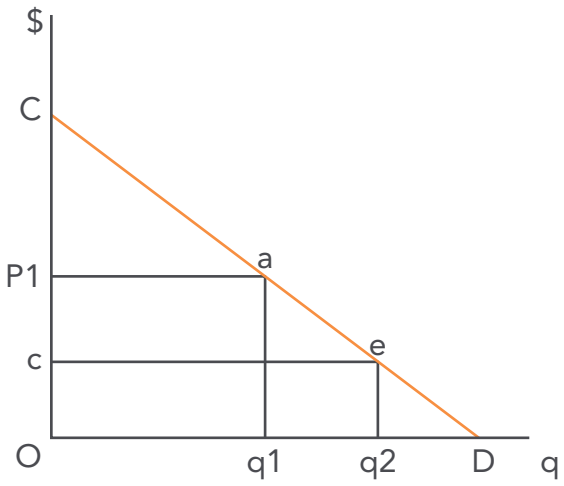


**Figure 6.4** Collusion

Assume there are two firms producing homogeneous products with no variable cost, marginal cost is zero. CD is the market demand curve. The first firm enters the market and maximizes profit by setting MR = 0, producing Oq1 which is half of OD and charging Op1. Firm 1 total revenue is Op1aq1 which is the largest rectangle under the demand curve. In this case firm 1 total revenue is equivalent of monopoly's revenue. Firm 2 faces segment of market demand curve and maximizes revenue by producing q1q2, half of q1D and charges Oc. The first firm notices that his price must be reduced to Oc or his sales drop to zero. At the price Oc both firms' total revenue is Oceq2 which can be shared between them. However, if they collude they can together produce Oq1 and charge Op1. If they collude their total revenue is the monopolist's revenue Op1aq1which is larger than Oceq2 and obviously each firm's share of total revenue will also be larger.

The Chamberlin's proposal is the theoretical basis for creation of a cartel where a group of firms agree upon a set price and sell according to the pre-determined quotas. The most famous cartel is OPEC (organization of petroleum exporting countries). The OPEC was successful in controlling oil prices in the 1970s where the demand for oil was strong. However when the demand is weak members attempt to enlarge their share of market by reducing their price. This development usually leads to the breakdown of the cartel.

**Game Theory**

The main behavioural feature of oligopoly is the **mutual recognition of interdependence among rival** firms in the industry. The sales of any one firm depend upon the firm's price as well as the prices charged by other firms. Before deciding to cut or increase its price, an oligopolist must try to predict how other firms will react and attempt to calculate the effects of those reactions on its own profits.

Recognised mutual interdependence leads economists to make use of a different form of analysis in studying oligopolistic behaviour. The modern economic method of analysing such interaction is known as **game theory** (derived from mathematics) where oligopolistic behaviour is modelled as a strategic game, with rules, strategies and payoffs.

Game theory has become a popular solution for an oligopoly market. The Theory of Games and Economic Behavior, published in 1944 by mathematician John von Neumann and economist Oskar Morgenstern, is acknowledged as the ground-breaking contribution to game theory. Another major contribution was developed by John Nash in 1950 in his PhD thesis in mathematics titled "non-Cooperative Games". In a game each player has a strategy. A dominant strategy is the best option that one player takes regardless of what the other one is doing. Nash equilibrium is the best strategy that one player takes given what other player is doing. A game may or may not have Nash equilibrium.

The pay-off matrix presents all of the possible pay-offs for all of the players in the game.

A simple example of game theory is Prisoners' dilemma.

Two suspects A and B are questioned in separate rooms in the police station for a crime that they have committed. A is told that if you confess and B doesn't confess you get a 2 years in prison and B gets 10 years. Each one of them knows that if they both confess, because of cooperation, each will receive 8 years. Because of lack of strong evidence, if both don't confess each will have 3 years in prison. However, they are separated and one doesn't know what the other is doing.

A dominant strategy in Game theory is the best option that one player takes regardless of what the other one is doing. The pay-off matrix of prisoners A and B is presented in Table 6.1

| A Strategies | B Strategies | |
|---|---|---|
| | Confess | No Confess |
| Confess | 8 years both | 10 years B, 2 years A |
| No Confess | 2 years B, 10 years A | 3 years both |

**Table 6.1** Prisoner Dilemma

In this case since each prisoner doesn't know what the other is doing, the dominant strategy for both is to confess. If both confess they get 8 years in prison. If they had not confessed it was only 3 years. But each one could not take the chance of the other confessing and only getting 2 years while they got 10.

Similarly in business suppose there are two manufacturers of laptop computer, Apple and IBM. They attempt to market their products through advertising on television or newspaper. The pay-off matrix of both firms is presented in Table 6.2.

| IBM Strategies | Apple Strategies | |
|---|---|---|
| | TV advertising | News Paper advertising |
| TV advertising | IBM profit $200 million<br>Apple profit $300 million | IBM profit $100 million<br>Apple profit $200 million |
| News Paper advertising | IBM profit $300 million<br>Apple profit $200 million | IBM profit $200 million<br>Apple profit $100 million |

**Table 6.2** IBM and Apple Pay-off Matrix

In the above example the dominant strategy for IBM is to engage in newspaper advertising because regardless of what Apple does, IBM makes more profit if engages in newspaper advertising. The dominant strategy for Apple is to emphasise TV advertising.

### The Effects of Oligopoly

The best examples of oligopolistic markets are the airline industry and car manufacturing. In both industries there are few firms producing similar but differentiated products. There are price competition and non-price competition in this market. To avoid consequences of price war generally non-price competition is preferable because reducing prices to sell more eventually could harm all of the producers in the market via a price-war.

Non price competition is mostly in form of advertising attempting to distinguish one product or service from the competing products. The entry to an oligopolistic market is not easy particularly if the concentration ratio (number of firms supplying large percentage of the market) is high and the incumbents erect entry barriers. Because of these characteristics oligopoly market prices are relatively stable. Generally the larger the number of firms in the market the lower is the price and the lower is the profit of individual producer. Often market participants attempt to keep the entry blocked by collusive arrangements.

### Strategic responses to potential competition, Some Practical Exercises

The analysis so far has concentrated on competitive behaviour within a competing group of firms. However, we have ignored the possibility of new competition from firms currently not producing the product but attracted by the profits being made. This potential or actual entry is an important structural determinant of prices and profits.

The following market situations may arise:

| | Profits of the Incumbents | Profits of the Entrants |
|---|---|---|
| Entry does not occur | High | Zero |
| Entry occurs; incumbents accommodate | Shared | Shared |
| Entry occurs; incumbents fight | Low or negative (due to price war) | Low or negative (due to price war) |

Firms adopt a long-term horizon which allows for changes in industrial structure through new entry. This means that short-run profit maximisation is often sacrificed in favour of maximising the present value of the future stream of profits.

One approach to avoiding market entry is to build barriers to entry, such as:

- product differentiation – creating brand loyalty; successful product innovation; patent protection
- absolute cost advantages
- economies of scale.

Another approach is for the incumbents to establish credible threats of losses for new entrants, thereby deterring or at least slowing down entry.

Let us look at limit pricing:

1. Draw a graph with price on the vertical axis and output on the horizontal. Have a downward sloping demand curve.
2. Draw in a horizontal straight line to represent the average cost of the incumbent firm. Entry barriers give an established firm a chance to raise price above its long-run average cost without attracting new competition.
3. Pick a price high above average cost on the demand curve.

The gap between this price and average cost multiplied by the quantity represents the firm's profit. However, too high a profit may induce new entry which means that the industry profit will need to be shared. So it is smarter to set a price (called the entry-forestalling price) that is lower than the entry-inducing price. Then you will deter entry and not have to share the profits with new competitors.

If there is a likelihood of new entry – if the market is potentially contestable – then this provides a restraining influence on price-setting.

Predatory pricing is where prices are set below (marginal) costs in order to kill off competitors and then raise prices to reap monopoly profits. The later extra profits more than make up for the early losses. Predatory pricing is only a section of the large domain of predatory actions which firms can use to suppress competition.

There are various non-price strategies that can also be used to deter entry – such as plant location and product differentiation.

**Pricing in practice**

There are a variety of models of pricing behaviour. We now look at empirical work to see how pricing decisions occur in practice. Numerous studies report that average cost pricing is widespread. The basic approach is to estimate unit costs and arrive at a price by adding a 'normal' or conventional margin to it, to allow for profit.

At times other pricing rules have been uncovered: various types of price discrimination, penetration pricing (i.e., setting prices that ensure a rapid increase in market share or sales to obtain economies of scale) or prices that are used to indicate product quality. In situations of intense competition, prices may be independent of cost conditions entirely.

Nevertheless, the most prevalent method of pricing is that based on average cost.

The profit margin is related to demand conditions. A 'rule of thumb' average cost pricing method with an adjustable margin can be seen as a useful device in an uncertain and complex market situation.

Studies also reveal that firms rarely attempt to maximise short-run profits. Pricing objectives involve long-run considerations of improving market share and achieving a target return on investment. Competitor's reactions and the threat of potential entry are also major concerns.

Since oligopoly lies between the two extremes of perfect competition and monopoly, the market outcome under oligopoly also varies between these two polar outcomes. The possible outcomes include:

- explicit collusion (cartel): oligopolistic firms agree to act like a monopoly (monopoly outcomes)
- implicit collusion: oligopolistic firms follow a price leader (a dominant firm) in changing prices
- non-collusive oligopoly: firms act independently (near competitive outcomes)
- non-price competition: firms compete using variables other than price and output.

**The Oligopoly Game**

An interactive game that students can play and thereby learn the essential features of oligopoly.

Here is the structure of the game:

Industry Demand P = 1500 – 5 Q

Marginal Revenue P = 1500 – 10 Q

Marginal Cost (each firm) P = 500 (so MC = ATC = 500)

Profit (each firm) Profit = P Qi – 500 Qi

Total Industry Output Q = sum of the Qi's

(25 is less than or equal to Qi which is less than or equal to 50)

Graphical Representation:

Draw a graph. P and Q on axes. Downward sloping D and MR lying below D (as in monopoly). Draw a horizontal line at P=500 to represent MC = ATC.

On the graph, when D intersects MC the Q = 200 and the P = 500 (the perfect competition outcome. Economic Profits = 0).

Then if MR = MC the Q = 100 and the P=1000 (the monopoly outcome. Economic profits = $50,000).

How to Proceed:

1. Spend 5 or so minutes talking generally about oligopoly. Note the interdependence of the rival firms. The strategic interaction in these markets. The incentives to collude. The threat of cheating. The outcome indeterminacy. How each firm is no longer in control of their own destinies as it depends crucially on how their rivals will react to whatever they do so you have to second guess what your rivals will do.

2. Divide the class into 4 groups of roughly equal numbers of students. These will be Firms 1, 2, 3 and 4. Get them to select a CEO. The CEO has to make the final decisions even though they can consult their underlings. Make sure each firm has a calculator. Get the 4 firms to congregate in the 4 corners of the room. The key is physical separation. Each firm must be distant enough from other firms that they are not part of that other firm's deliberations and discussions. So keep them well apart.
3. On the board give them the structure of the game and graphical representation (above). I usually put this up already for them before the tutorial starts. Also put up on the board a spread sheet with 8 columns.

| | *Round 1* | | *Round 2* | | *Round 3* | | |
|---|---|---|---|---|---|---|---|
| | *Qi* | *Profiti* | *Qi* | *Profiti* | *Qi* | *Profiti* | *Total Profit* |
| *Firm 1* | | | | | | | |
| *Firm 2* | | | | | | | |
| *Firm 3* | | | | | | | |
| *Firm 4* | | | | | | | |
| | *Sum Qi =* | | *Sum Qi =* | | *Sum Qi =* | | |
| | *P =* | | *P =* | | *P =* | | |
| | *Industry Profit =* | | *Industry Profit =* | | *Industry Profit =* | | |

4. Stress that the objective of each firm is to maximize profits. There are 3 rounds of the game (or 3 years if you like) and the overall winner is the firm that gets the most profits over the 3 years.
5. Tell them that there are no cost differences between the firms. There is no product differentiation. They are all producing the same product. There only decision variable is to decide what output they will produce and they are restricted to a Qi that is between and including 25 and 50.
6. Tell them that Round 1 begins now. Let the firms have a brief discussion among their employees as to what output to produce BUT they should be careful that they do not let their rival firms know what they have decided. Then go to each firm and ask them on a piece of paper to write their firm number and there output level. Fold the paper so the other firms can not see it. Once you have collected all 4 firm outputs then share the results.

   Let's say they all select 50 (the max. output allowable). Fill in the table. Each Qi is 50. The total industry output is 200. The price from the demand equation is then 500. Each firms profits (ask the CEOs to work this out for themselves) is then $500 \times 50 - 500 \times 50 = 0$. So they have replicated a perfectly competitive situation.

   It is more likely that they provide different outputs. Let's say it is 35, 50, 25, 33. Then industry output is 143. Price is $785 and the respective profits are 9975, 14250, 7125, 9405. (You need the calculators to work out the individual firm profits).

   I usually add to the fun by sacking the CEO of the worst-performing firm (in this case – firm 3 that only made 7125) and lavishing praise and perks on the best performing firm (in this case firm 2).

7. Then it is off to Round 2 and Round 3 and the eventual winner. After each round take an opportunity to analyse the results. Note that the most profitable outcome for the industry as a whole is for them all to produce 25 and equally share industry profits of $50,000 (so that is $12,500 each). But as we saw with firm 2 if one cheats (produces more) they win while the only firm that opted for 25 (firm 3) was the sacrificial lamb. Keep repeating to the students that they could be earning $12,500 each (i.e. reinforce the pressures to collude).

8. You will get a variety of results and opportunities to analyse and discuss each year's results. You may get a PC outcome with zero profits. Or they may collude (this is unusual – they are usually too timid). But there will be a lot of cheating and revenge and sacrificial lambs. I spice it up with all female firms and other ways to segregate the firm's members. I threaten the firms that continually perform poorly with being taken over by corporate raiders. I tell them that this is just for fun but in the real world it is real money, prestige & status, collusion, successful CEOs are head-hunted by other firms, and that the more profits the industry makes the less is consumer surplus and the overall welfare of consumers. Mention the regulatory agencies and the penalties for collusion. But don't do anything fancy until rounds 2 or 3 (other than sacking very poor CEOs) so they understand what is going on first.
9. After 3 rounds declare a winner and sum up their experiences.
   We find this a very popular game and students learn more about oligopoly from it than they do from a text or lectures. It becomes real for them. Sometimes the only thing they remember about Micro I ten years down the road is the oligopoly game they played! Have fun.

Feedback by Instructors:

Instructor 1:
My students really seemed to enjoy the oligopoly game in tutes this week.

They had a lot of fun and I definitely think they learnt a lot. They were a bit confused at first about how to work out equilibrium price and quantity in the market under a perfect competition scenario and a monopoly scenario – perhaps next time it would help to give them the game ahead of time so that they can make sure they understand the mechanics of the game before hand?

Instructor 2:
The oligopoly game was useful as a basis for discussion of market behaviour of firms in oligopoly, however some students understood the principles behind the game whereas others didn't. I expect that reflects the different abilities and interests of the students anyway.

Instructor 3:
What it shows clearly is how rapidly the perfectly competitive solution comes in, and how strong the incentive to collude is. One student remarked that it would be more interesting to set a level of profit to be achieved. That was interesting in that it introduced a different objective (satisficing) into the discussion. To me it was interesting watching the different personalities shine through. My own view is that the problem should be made slightly more complicated by introducing one extra variable. But overall they had fun.

Instructor 4:
I have run the game in 3 classes. It was fine in general and students had fun as well, at least they seemed to. It was a useful way to introduce the concept of price war and collusion in oligopoly context.

Instructor 5:
The response to the game has been fantastic.

Instructor 6:
I think overall the students really enjoyed the oligopoly game, most students understood what was going on by the end of the second round.

I had to supplement the oligopoly game by getting volunteers up to play basic prisoners dilemmas, along with discussing what could happen if the prisoners dilemma is played repeatedly.

Instructor 7:
I think my classes enjoyed the game. They were taking it beyond the oligopoly area and into different areas. One team turned to bribery. A lot of cheating. Overall they opened up and had a great time.

Instructor 8:
The oligopoly game went well generally in the tutorials. For the first round, most teams just randomly selected a number. The game then progressed with each team increasing their production output in round 2 and round 3, with all firms choosing 50 in round 3 in most tutorials. Students were able to get the idea that in a oligopoly market structure there is a mutual interdependence and that it is very difficult to collude. A few students suggested that the firm which performed the worst in each round should be eliminated from the competition to make it more interesting.

**References**

Chamberlin, E. The Theory of Monopolistic Competition, Harvard University Press, Cambridge Massachusetts, 1933.

Cournot, A., Researches sur les Principles Mathematiques de la Theories des Riches, Translated by National Bacon, New York: Macmillion Co., 1897.

Mansfield, E. Microeconomics: Theory / Application, Eighth Edition, W.W. Norton & Company, New York, 1994.

# 7 INPUT MARKETS

Inputs such labour, capital and land can be purchased under different market structures. Here only the market for labour is discussed, similar analysis is applicable to other inputs.

Two cases are considered in this chapter:

1. Inputs are purchased in a perfectly competitive market and products are also sold under perfect competition.
2. Both inputs and output are purchased and sold in less than perfectly competitive markets.

**1. Perfect competition in both input and output markets**

The market price for an individual producer in a perfectly competitive market is constant. Also the wage rate for an individual buyer is constant when labour is purchased under perfect competition.

Every unit of labour hired adds the value of its marginal product VMP, to the revenue of the producer and adds to the cost equal to the wage rate, w. Labour will be hired up to the point where profit is maximized. Labour employment stops when VMP = w.

$$VMP = MP \times P$$

where MP and P are marginal product of labour and market price respectively. VMP is change in revenue when one unit of labour is added. The derivation of demand for labour based on profit maximization and diminishing marginal product of labour is presented in Figure 7.1.

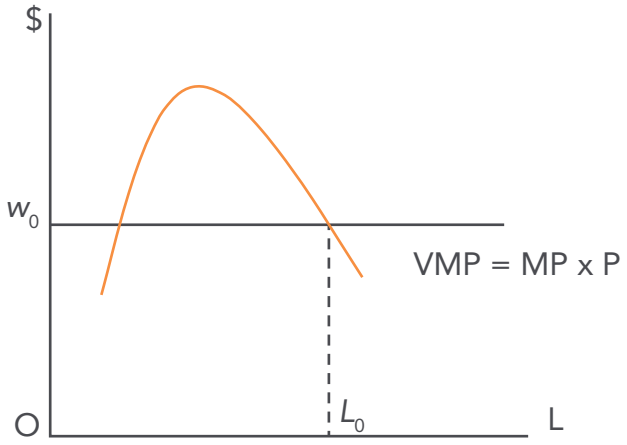


**Figure 7.1** Labour Demand Curve

In Figure 7.1 assume that the wage rate is $Ow_0$. Labour is hired as long as VMP > $Ow_0$. Profit maximizing employment is $OL_0$ where VMP > $Ow_0$ and the slope of VMP is negative. Before $OL_0$ a unit of labour adds more to the revenue than to the cost, VMP > $Ow_0$causing profit to rise (if there is profit). Profit falls after $OL_0$ where VMP > $Ow_0$. If profit falls after $OL_0$ and rises before $OL_0$, it must be at maximum at $OL_0$ where diminishing marginal productivity applies. Accordingly, the necessary condition for profit maximization in employment is the equality of VMP and the wage rate. The sufficient condition is diminishing marginal productivity of labour.

The demand for labour is the downward sloping segment of the VMP curve. Every point on that section of the VMP shows the level of employment at different level of nominal wage.[6]

The market demand for labour can be derived from horizontal summation of all of purchasers' VMP curves. The market demand is also downward sloping indicating that employment rises as wages fall.

The supply of labour is based on substitution of work for leisure. As wages rise people tend to work more and spend less time on leisure. An individual labour supply curve is upward sloping. The market supply curve is derived from horizontal summation of all of the individual's supply curves.

The wage rate and employment when perfect completion exists in input and product markets is determined at the intersection of market demand for and market supply of labour, Figure 7.2.

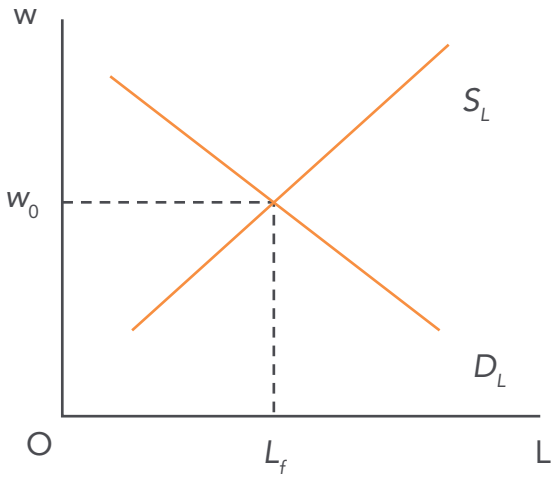


**Figure 7.2** the Labour Market

At $Ow_0$ the labour market is at full employment where $S_L = D_L$. At any wage rate above the equilibrium $S_L > D_L$ leading to unemployment.

### Rent and Quasi Rent

The concept of rent applies to the payment to an input which is fixed in quantity. For example payments to land which is fixed in quantity, payments to a movie star or a sports – men who are unique and rare in their professions. This rent is different than the rent which is paid for the use of other means of production.

Since the supply of an input is fixed the rent or the price of an input is determined by the demand. The payment of rent is presented in Figure 7.3.

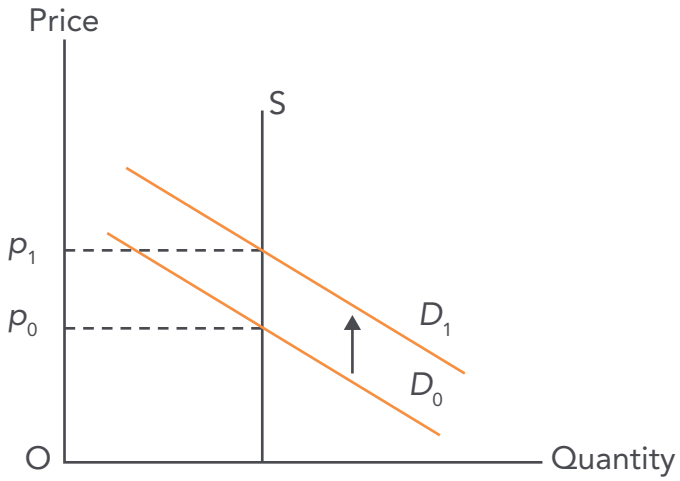


**Figure 7.3** The Concept of Rent

In Figure 7.3 the rent is $op_0$ when the demand is $D_0$ and rises to $op_1$ when the demand rises to $D_1$. The rent is entirely determined by the demand for the input.

Quasi rent is the return to the fixed factor of production which is temporarily fixed in the short run. The firm has to pay the variable cost to keep their employment. The residual which is left is the quasi rent. This is illustrated in Figure 7.4.

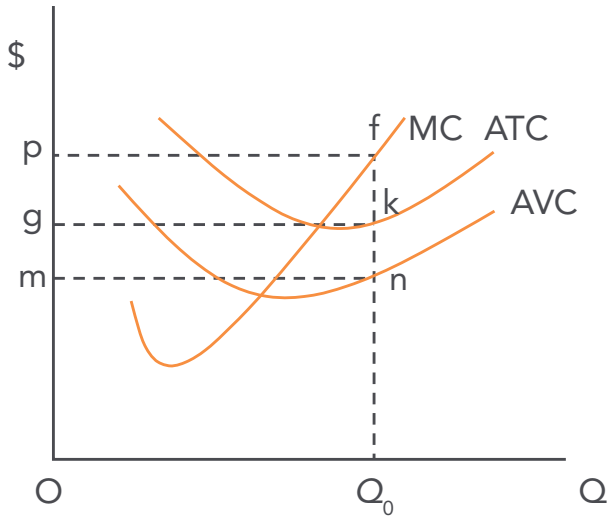


**Figure 7.4** Quasi Rent

Suppose the market price is op. The profit maximizing output is $OQ_0$. The payment to the variable inputs required to keep them in production is Omn$Q_0$. The residual mpfn is the quasi rent. The quasi rent is different than the fixed cost which is mgkn. The quasi rent is the difference between total revenue and total variable cost. The quasi rent could be equal to zero if the market price is equal to the AVC.

### Efficiency Wages

It is often observed that workers, if paid low wages which is paid in other enterprises, do not provide their maximum work effort. Because it is difficult to measure low performance, managers are better-off to pay efficiency wages (higher wages) in order to create incentives for workers to increase their productivity. The opportunity cost of an efficiency wage is high because those wages are not paid in other enterprises. Figure 7.2 illustrates the effects of minimum wage and the efficiency wage on employment.

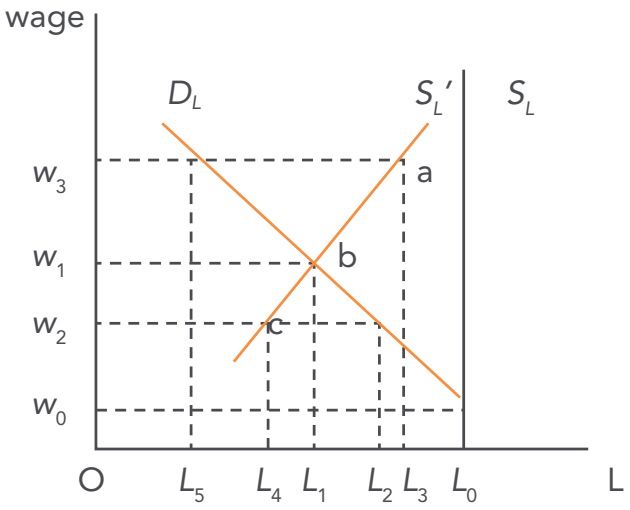


**Figure 7.5** Efficiency Wages

At $Ow_0$ demand and supply of labour are equal to $OL_0$. $Ow_0$ is the wage rate, which is paid in other enterprises. At this wage rate workers have little incentive to work hard because the opportunity cost of losing their job is low. They can always find employment in other places with the same wage rate. In order for workers to perform better, the firm has to pay them higher wages than $Ow_0$. At the higher wages workers work harder because if they lose their job they have to accept other jobs with substantially lower wages. In Figure 7.5 the $S_L$' curve is the demand for labour when efficiency wage is paid, points such as a, b and c. $S_L$.' is inversely related to the level of unemployment. For example, at $Ow_2$ unemployment is $OL_0 - OL_4$. At $Ow_3$ unemployment is $OL_0 - OL_3$ which is smaller than unemployment at $Ow_2$. The equilibrium efficiency wage is at $Ow_1$ where the demand for labour curve intersects the efficiency wage demand for labour curve. It can be shown that any other wage rate will return to $Ow_1$.

## 2. Imperfect Competition in Input and Output Markets

AR and MR curves are different when monopolistic condition exists in the product market. A single purchaser of inputs is called monopsony. Wages rise when a monopsonist purchases more labour. In this case the wage rate is always greater than the marginal cost of labour (MCL), when the average is rising marginal is greater than average. In the product market a unit of labour adds MRP = MR × MP to the revenue since AR is different than MR. MRP is downward sloping because both MP and MR are downward sloping. Profit is maximized at $OL_0$ where MRP = MCL. However, the wage paid to the labour is less than MRP. In this case the monopsonist exploits the labour by paying discriminatory wages. Employment and wage rate when monopoly exists in both markets is presented in Figure 7.6.

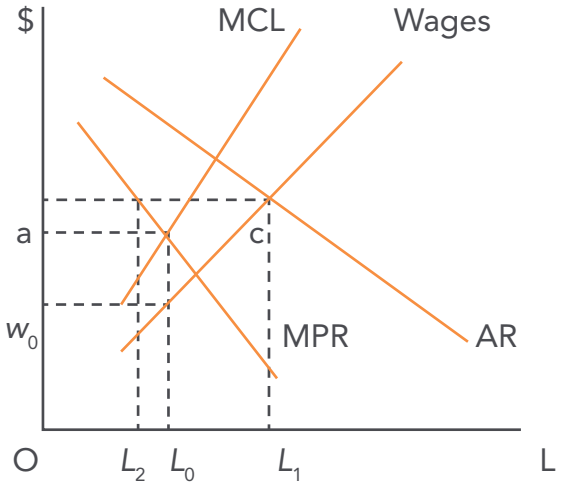


**Figure 7.6** Monopsony

In Figure 7.6 profit maximizing employment is $OL_0$ and the wage rate is $Ow_0$. The distance $aw_0$ is the extent of discrimination where labour adds more to producer's revenue than is paid.

## Labour Unions

Labour unions attempt to improve working conditions of labourers. They can bargain for wages to prevent exploitation, maintain living standard of their members or secure their employment. Union membership is associated with the largest boost to wages at the lower percentiles of the income distribution but by 2007 in the US there was a two-thirds reduction in private sector union coverage compared to 1973.

In Figure 7.6 the competitive wage rate and employment is at point c where supply of labour is equal to AR curve. If they are successful in their bargaining, this wage rate increases both employment and wages. The most likely outcome is to set the minimum wage rate at cL1 which reduces employment. The effect of setting minimum wage at point c on employment is smaller the lower is the elasticity of demand for labour. The elasticity of demand for labour is % Δ L / % Δ w. When the elasticity of demand for labour is less than unit y, % Δ L < % Δ w causing total labour force income to rise as a result of an increase in the wage rate.

## Welfare Effects of Monopsony

In Figure 7.6 it was shown that a monopsonist hires less labour and pays lower wage compared to a perfectly competitive input market. This behaviour generates deadweight loss to the society. The effects of the monopsony are presented in figure 7.7.

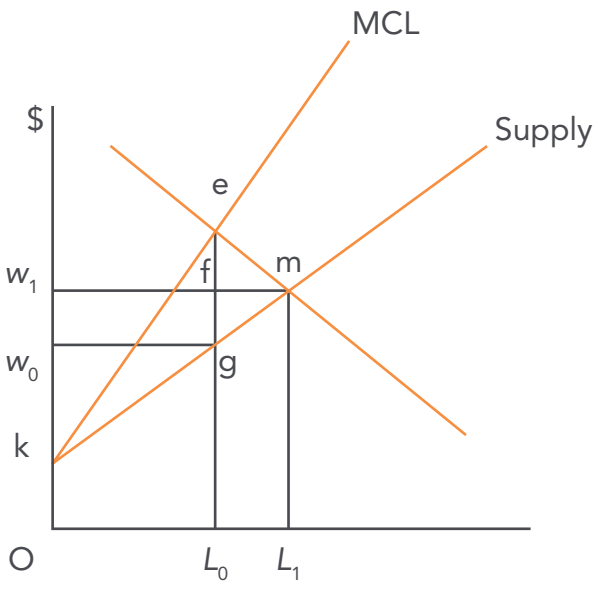


**Figure 7.7** Welfare Effects of Monopsony

The monopsonist hires $OL_0$ and pays O$w_0$. The employment and the wage rate would have been $OL_1$ and pays O$w_1$ if the input market was perfectly competitive. The labours' surplus is km $w_1$ if the input market was competitive. Under monopsony the labours' surplus is k$w_0$g. The area of the rectangle $w_0w_1fg$ is transferred from the labours' to the producers as they pay lower ages and hire less labour. The triangle egm is the deadweight loss of the imperfect competition in the input market.

**The Distribution of Income**

Factor pricing outcomes have clear implications for the distribution of income and income inequality has increased in more than three-quarters of OECD countries from the mid-1980s to mid-2000s. The postwar boom – the so-called 'Golden Age' of Capitalism (or the Long Boom) – that lasted from between 1947 and 1973 was characterized by rapid economic growth that was equally distributed among income classes. However, in the decades since then many economies have underperformed and failed to deliver broad-based growth.

The disparities in the U.S are particularly stark. The top 1 percent of American households accounted for only 8.9 percent of income in 1976, but this share grew to 23.5 percent of the total income generated in the United States in 2007. Of every dollar of real income growth that was generated between 1976 and 2007, 58 cents went to the top 1 percent of households. In 2007 the hedge fund manager John Paulson earned $3.7 billion, about 74,000 times the median household income in the United States.

What were the causes of the increased U.S income inequality? Technological advance has reduced the demand for certain lower-skilled occupations while increasing the demands for highly educated and trained workers. Moreover, an influx of lower-skilled migrants and cheap labour-intensive imports from places like China has placed pressure on the wages of lower income Americans. However, while skills-biased technical change, immigration and trade are common explanations at best they point to a rising gap between less-educated and more-educated workers but that is only part of the inequality increase – even the college educated have seen wage gains lag behind productivity. It does not explain the excessive gains at the very top of the income distribution.

A major issue of discussion is what has happened to middle-class America? Some studies report that while US GDP per capita increased by 10 percent between 2000 and 2008 – median household income was 4% lower. The median income of males in their thirties today is reported to be lower than it was three decades ago. There is heated debate on whether American median income has risen or fallen since the early 1970s. Studies suggest that it has either risen modestly or actually declined depending on whether you are just looking at males or households, whether it is adjusted for inflation and hours worked, and the time period studied. Looking at families it has risen by about 16 percent from 1973 to 2005 but the increase has mainly reflected longer working hours through greater female labour force participation. Since 1980 median family income has risen only about 0.7 percent a year. For men, median inflation-adjusted earnings have fallen slightly. If the gains in productivity had been evenly shared across the workforce, the typical worker's income would be 35 percent higher now than it was in the early seventies. While the top 10 percent of the wage distribution saw spectacular gains, the rest of the distribution achieved far less than the 35 percent average. Average income has gone up substantially, as has labour productivity, but not median income.

Postwar America was a middle-class society with narrow income gaps, and a far more equal distribution of income and wealth. Ownership of wealth was less concentrated. There was no extreme inequality. Most people lived recognizably similar and remarkably decent material lives. It was a 'golden age' that lasted until the 1980s. Thereafter, inequality increased as the welfare state was rolled back, unions became less influential and tax rates on the wealthy were reduced – there were dramatic drops in the top marginal tax rate. Between 1979 and 2006 the top tax rate on earned income was cut in half, the tax rate on capital gains was cut almost as much and the tax rate on corporate profits was cut by a quarter. The result was runaway growth of executive pay while at the other end the minimum wage in 1966, in real terms, was higher than it is today. In 1966 the typical male in his thirties was earning as much as his modern equivalent.

**Taxation for Income Redistribution?**

In the earlier parts of this chapter we talked of wages and made no distinction between before and after tax earnings. How might the tax system be used to redistribute income? Most countries have a progressive income tax system that taxes extra income at a higher marginal rate of taxation. What are the consequences for work effort? Let us take a simple example.

Assume we only have 2 people in an economy – Person A and Person B.

Their before-tax income is:

Person A = $100,000

Person B = $20,000

National Income = $120,000

This is an inequitable situation, Person A earns 5 times more than Person B or 83% of the country's income goes to Person A and Person B only gets a paltry 17%.

In order to restore "fairness" we introduce a progressive income tax. The tax free threshold is $60,000 so Person B pays no tax. Every dollar earned above $60,000 is taxed at 100% (so that is the marginal tax rate which in this case is also the average tax rate). So Person A is taxed $40,000 and that tax revenue is redistributed to Person B as a welfare payment. The end result is after tax income of:

Person A = $60,000

Person B = $60,000

National income = $120,000

We have a perfectly equal society and no loss in national income and so this is a case of no equity/efficiency trade-off.

But we have made some of the assumptions fairly extreme – the 100% tax rate – so we can clearly see the problems.

As Arthur Okun long ago said the redistribution of income from rich to poor takes place in a 'leaky bucket'. If the bucket had no holes the full $40,000 goes to the poor person. But there are 'leaks' and so only some of that $40,000 ends up in his/her pocket.

The first leakage is the cost of tax administration and welfare services. The tax system needs to be administered and tax compliance monitored and then social services need to allocate resources to the poor, after appropriate paperwork is submitted and processed, which may not involve a complete monetary transfer but 'help in kind' (job retraining skills and personal budgeting advice) and directives about how the money is to be spent (a certain amount must be housing specific or for dependent children or for 'healthy' foods etc). So the costs of the redistributive tax system and the social service (income support) agencies can be significant. That could easily eat up $5,000 of that $40,000 in public service salaries alone, leaving only $35,000 available for the poor person.

A second leakage may come from the rich person. All of his income above $60,000 is taxed away. Why should he work hard and be entrepreneurial and put all that effort in if all that extra income is taken away and given to a perhaps 'undeserving' other person? The rich person may just work sufficient hours to generate $60,000 and use the rest of the time for leisure purposes. Alternatively, he may hire an expensive tax account or tax lawyer (more leakages) to engage in tax avoidance or even tax evasion. All the available tax loopholes and concessions will be exploited or all income above $60,000 may be transferred to a tax haven. The end result is that far less than $100,000 will be reported as taxable income.

The third leakage may come from the poor person. They are working full time for $20,000 and getting another $40,000 from welfare. They are doing just as well as the rich person. They can slacken off, claim a bad back and go on disability support. They were 'surviving' on $20,000 and now they can enjoy leisure time and sit back and even without work get $40,000 – off to the beach they go!

Worst case scenario of reported income before tax:

Person A = $60,000

Person B = 0

National Income = $60,000

We still have a terribly unequal income distribution but National Income has declined by 50%. That is a very high cost (a very high trade-off) for an attempted income redistribution program. What do you do now? Remove the income tax threshold and impose a 50% tax rate so that each person ends up with $30,000?

Obviously we have used extreme assumptions to clearly demonstrate the issues that can arise. That makes the potential problems very clear. More realistic assumptions (say a top marginal rate of 45% for incomes over $180K) would modify some of the implications. But the basic ideas remain the same.

**Factor returns and education**

What are some of the distinctive features of the increased inequality in the U.S? Income redistribution has hurt minorities disproportionately more. The African-American poverty rate is triple the white rate and the unemployment rate is double. There is also a strong sense of unfairness, injustice and deprivation on the part of African-Americans in that they feel they have been dealt a bad hand in terms of resources of their parents and resources of their neighbourhoods. They may legitimately complain that they lack the skills and assets to compete effectively in the current labour market partly because of the low quality of schooling they have received. It might be noted that in the U.S only 34 percent of youths from households in the bottom quintile of the income distribution go to college, whereas 79 percent of those from the top quintile do. Only 11 percent of the bottom quintile graduate while 53 percent of the top quintile do. Middle-class America is also worried about their children's future and concerned that they are not getting the quality education they need to compete in the labour market.

Empirical studies support more funding for early childhood support in terms of better nutrition, health care, and early education initiatives – preschool programs and day-care centres. They support further financial aid to encourage participation in tertiary education with studies suggesting a $US1000 student subsidy increases college attendance rates by roughly 4 percentage points.[7]

# 8 EXTERNALITY, PUBLIC GOODS AND GOVERNMENT REGULATIONS

Externality is any costs or benefits associated with consumption or production that are not included in the market price. In other words, the private market system doesn't indicate the socially desirable levels of consumption and production of goods and services with externalities.

Examples of externality in consumption are smoking, loud disturbing music and driving dangerously. In production, externalities are air and water pollution, planting flowers with nice aroma in public roads, research results benefiting others. Externality in consumption occurs when the private demand for a good or service diverges from the social demand. In production, externalities refer to the divergence of private production from the socially desirable level of production.

External costs in consumption and production are presented in Figures 8.1a and 8.1b.

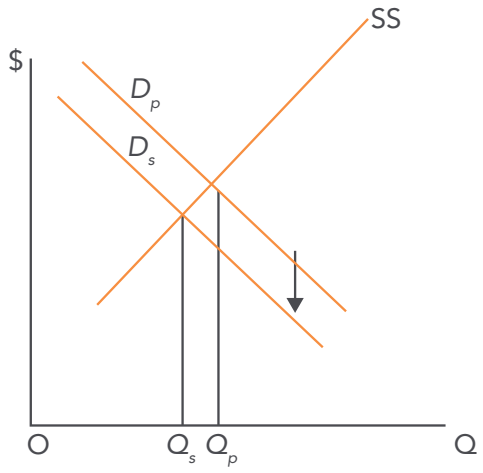


**Figure 8.1a** Externality in consumption

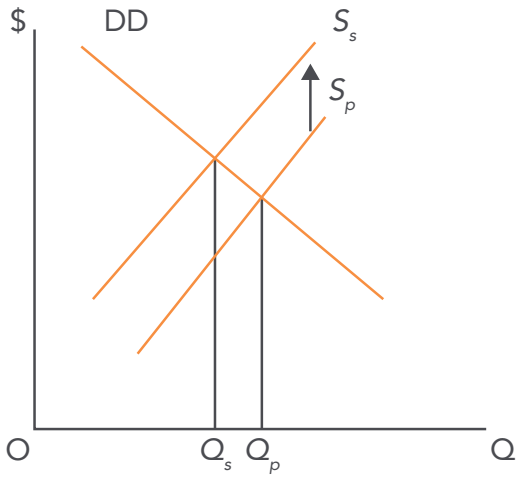


**Figure 8.1b** Externality in Production

In Figure 8.1a $D_s$ *and* $D_p$ are social and private demand of a good with negative externality in consumption (smoking). $Q_s$ *and* $Q_p$ show that the social consumption of this good is less than private consumption. This indicates that, because of extra cost to the society, the society prefers less consumption of this particular good.

In Figure 8.1b negative externality in production of a good (pollution) is presented. The society prefers less production of this good relative to what is produced by the private sector.

In case of goods and services with positive externality, the socially desirable quantities of those goods and services are greater than the private quantities.

### Government Regulations

Externalities can be internalized by government regulations. Taxes and subsidies are used to move the private quantities of goods and services with external costs and benefits to the socially desirable quantities. In Figure 8.1a taxes on consumption of a good with external cost (sales taxes on cigarettes) moves the SS curve to the left causing the equilibrium price to rise and the quantity to move closer to the socially desirable level. Higher prices on cigarettes discourage smoking. In Figure 8.1b Taxes on production of a good with negative externality moves the SS curve to the left causing an increase in price and a reduction in quantity produced. The effects of subsidies are opposite of the effects of taxes. Subsidies are paid out for consumption and production of goods and services with external benefits. Subsidies tend to encourage consumption and production of goods and services and lower the equilibrium price. Subsidies paid to researchers, farmers and exporters are examples of government regulations to provide incentives for consumers and producers to increase quantities consumed and produced.

### Environmental Pollution

Environment pollution in form of smogs, dumping rubbish in rivers etc, impose cost on the society. Removal of pollution also bears cost. Generally people think that an economically efficient level of pollution is zero level. However, for determining the efficient level both costs of pollution and costs of pollution removals have to be considered. Consider Figure 8.1, Mansfield (1994), where cost of pollution, cost of pollution removals and the sum of the two are presented. The horizontal axis shows the quantity of pollution emitted.

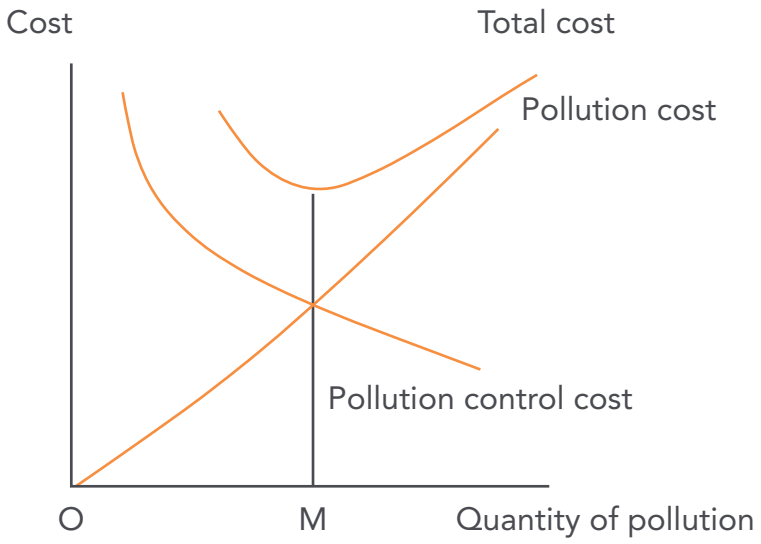


**Figure 8.2** Efficient quantity of Pollution

In Figure 8.1 the cost of pollution rises as the quantity of pollution increases, the upward sloping curve. The pollution control cost is inversely related to the quantity of pollution, this is the downward sloping curve in the graph. The efficient quantity of pollution is OM where the total cost is at minimum. At OM the cost of pollution is equal to the cost of pollution control. To the left of OM an additional unit of pollution reduces the cost of pollution control by more than the cost of pollution causing the total cost to decline. To the right of OM an additional unit of pollution increases the cost of pollution more than it reduces the cost of pollution control causing total cost of pollution to rise. If total cost is falling before OM and rising after OM, it must be at minimum at OM. It has been shown that the efficient level of pollution is not zero pollution.

In practice there are several pollution control methods. Two popular methods are an effluent fee and transferable emission permits (ETP). Effluent fee is like imposing tax on the polluter in Figure 8.1b. The fee increase the marginal cost of producing an additional unit of output, causing the supply curve to shift up. At the new equilibrium, lower production reduces the quantity of environmental pollution.

ETP are licenses sold to firms allow them to pollute a fixed amount (an efficient quantity). The total quantity of pollution by all of the polluters is fixed. These permits may be bought and sold. Those that find it profitable to pollute buy shares from those who do not need to pollute much. The exchange of permits leave the total amount of pollution unchanged. The total quantity of pollution remains fixed at the efficient level.

**Property Rights: Coase Theorem**

R. Coase in 1960 argued that under special conditions, despite existence of externalities a private competitive system can provide an efficient allocation of resources. The special conditions are:

a. The existence of well-defined property rights,
b. Negotiation between parties with minimal transactions costs.
c. Small number of parties involved.

The Coase theorem can best be explained by an example. Suppose there are two neighbours living in two separate houses with well-defined property rights. Both are entitled to having clean backyards. One of the neighbours cuts his trees and dumps garden refuse in the other courtyard creating extra cleaning costs for the other person. According to Coase theorem, they can negotiate to remove this extra external cost. If the dumping neighbour pays the second person for cleaning his court yard, the externality is removed and the marginal private cost of cutting trees will be equal to the social marginal cost.

The Coase theorem works without government intervention provided that the above conditions hold. The transactions costs of negotiation become high if the number of negotiators is large. In this case negotiation will be too costly and impractical.

**Public Goods**

Public goods such as national defence, roads and national health system have two important features; Non rivalry and non exclusivity. Non rivalry means that the marginal cost of consuming one more unit of a public good is zero. Non exclusivity implies that once goods are produced everyone can use them.

Public goods are subject to the free rider problem, that is; those who don't pay for them may also consume them. National defence is the extreme example of public goods. Once it is established everyone can use the protection and security of national defence. However, there are public goods that can be privatized by imposing certain barrier to consumption. For example, roads, bridges and airports are public goods that can be exclusive by imposing tolls. One cannot use some of the roads and bridges without paying tolls. In this case consumption of a public good is restricted to those who can afford to pay for it.

The efficient allocation of public goods is determined at the intersection of demand and the supply of public goods. However, derivation of the total demand for a public good is different from the derivation of the total demand for a private good. In case of public goods, the total demand curve is vertical (not horizontal) summation of each individual's demand curves. The efficient quantity of a public good is presented in Figure 8.3.

Suppose there are two individuals A and B consuming a public good. The total demand curve DD, is the vertical summation of A and B demand curves.

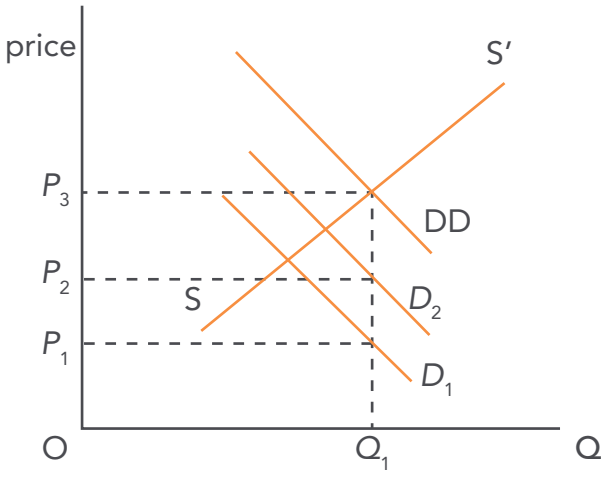


**Figure 8.3** Efficient Quantity of a Public Good

In Figure 8.3 SS' is the supply curve and the efficient quantity of public good is $OQ_1$. The total price paid for the public good is $OP_3$ which is the sum of $OP_1$+ $OP_2$. $OQ_1$ is the efficient quantity of public good because at that quantity the marginal social cost of the public good is equal to its marginal social benefit. Often public goods are subject to free rider. Consumers know that their consumption creates no extra cost; they find it worthwhile to consume the good without paying for it.

Public goods may be provided by the private sector. For example a few households located in an isolated place may pull their resources together and build a road that is used by all of them. However, in large societies public goods will not be produced without government intervention.

### Cost and Benefit Analysis (CBA)

The CBA is a method for evaluation of government projects. The projects are ranked in order of their benefit cost ratios. If there is no budget limitation (very rarely) those with ratios higher than unity are selected. However, when a budget constraint exists, those projects with higher benefit cost ratios will be selected provided that their total cost does not exceed the budget. A simple example of CBA is presented in table 8.1.

| Projects | Benefits $ | Costs$ | Benefit/Cost ratios |
|---|---|---|---|
| A | 150,000 | 120,000 | 1.25 |
| B | 40,000 | 60,000 | 0.66 |
| C | 550,000 | 400,000 | 1.375 |
| D | 800,000 | 300,000 | 2.5 |
| E | 900,000 | 600,000 | 1.5 |

**Table 8.1** Cost Benefit Analysis

In Table 8.1 all of the projects except B will be selected if there was no budget constraint. However, with the presence of $1 million budget, projects D and E are feasible. If E is not selected then projects A, C and D can be chosen.

### Some examples of government regulation that may improve community welfare

*Financial Markets*

In a well-functioning market economy, markets are supposed to provide the incentives that lead individuals to do what is in society's interest. The Global Financial Crisis of 2008 and 2009 has clearly indicated that America's financial markets had failed to perform their essential societal functions of managing risk, allocating capital and mobilizing saving while keeping transaction costs low. Financial markets did not allocate capital to its most productive use where the returns to society were highest. Instead they misallocated capital, engaged in excessive risk-taking and lending to those that could not repay, leading to a situation where private rewards were unrelated to social returns. Indeed, banks performed so poorly at credit assessment and mortgage design that they have put the entire economy at risk.

Many observers lay the main responsibility for the global financial crisis on financial markets and institutions for reckless lending leading to a major private-sector led misallocation of financial resources. The financial sector, circumvented regulations and accounting standards, created excessive risk, encouraged excess debt, excessively focused on short-term returns and received exorbitant executive pay that had little relation to performance. Financial engineering led to products so complex that those seeling and buying them did not understand the risk implications. Poor corporate governance, clear conflicts of interest, perverse incentives and externalities meant that private rewards were not well aligned with social objectives and the financial sector redistributed wealth rather than created new wealth and as such provided little social benefit. The Government has had to repeatedly save financial markets from their own mistakes.

*The problem of Obesity*

Medical theories lead to medical solutions to obesity – pills that alter metabolisms, molecules that affect fat absorption or surgical procedures. Psychologists and sociologists might look at cultural explanations. Eating habits are influenced by social interactions; we tend to eat more in a group setting compared to when we are alone. There are also national differences. People in the United States like larger meals than do typical French consumers and the French take longer to finish those smaller meals. However, economists tend to analyse health problems in terms of price signals, the responsiveness of consumers to incentives and the overall influences on demand and supply for products and services.

It is clear that genetics contributes to diabetes and obesity but the rapid increase in the incidence of these health issues in just a few recent decades indicates that other factors are at work. We need to look at the price of healthy and less-healthy food alternatives, the changing nature of demand for processed food and its supply, and particularly how it has been impacted by time-poor households. In earlier times, when labour participation rates of married women were low, the traditional bread-winner came home to a home cooked meal that required considerable preparation. With increased female participation in the workforce, families are often time poor – the demands of family and work drain parents of the energy to exercise and to prepare elaborate meals. Often if the choice is to spend an hour cooking every night or simply pick up some KFC or pizza on the way home from work, the easier alternative is selected. A large proportion of the obesity epidemic is due to increased purchase of food that is prepared outside the home.

A further factor is the change in food prices over the last few decades. The food industry has developed more efficient ways to manufacture cheap, energy-dense foods and perfected ways of storing and packaging these foods that reduce the time cost of eating them. People eat far more fast food and snack food than they used to. Processed foods are far quicker to prepare than healthier food. The relatively higher prices of healthy foods – like fruit, vegetables, and whole grain breads – create incentives for families to purchase less healthy foods. Price signals and advertising strategies encourage us to move away from the produce section to the processed food section. With the lower prices, it is then no surprise to see that the incidence of obesity is greater in low income families. If you combine the sometimes exhausting pressures of looking after a young family, worrying about financial problems, bad work environments, traffic congestion or neighbourhood crime issues, then we can partly explain the greater rise in obesity in poor families over the last few decades. These groups are both time-poor and budget-constrained.

Both the demand for, and supply of, processed foods has expanded considerably over the last few decades and this leads unambiguously to an increase in the purchase and consumption of foods that accumulate calories. At the same time, we are exercising less. In terms of physical fitness, technological change has reduced the number of calories expended by labourers. In earlier times workers were paid to exercise (work required a lot of physical effort) but now they are operating highly mechanized machines or sitting in front of computer screens all day. If they want to exercise they have to pay for it (e.g. join a gym) so the cost of exercise has gone up. Also, when they come home after work, they have far more choices about how they will spend their time – they can exercise or get on the internet and surf the web or get into the various social media sites or watch Foxtel or play interactive computer games. The alternatives (opportunity cost) to exercise have become far more attractive.

So what can be done? In theory losing weight should be simple, just burn more calories that you consume. But only five percent of people who try to lose weight manage to do so in a sustained manner. Self-control problems are very important and often the benefits of immediate gratification count more heavily than the longer term harmful consequences (clogged arteries and failed kidneys). The pain of forgoing delicious meals is perceived to be greater than the benefits of reducing one's calorie intake. We need informed consumers, knowledgeable about the risks and benefits of what they are eating, so there is certainly justification for mandating information about calorie content. We can also use price incentives – subsidize healthy foods and tax unhealthy ones. We could even subsidize gym membership. Yet would we anticipate significant shifts in buying behaviour? That depends on demand responsiveness to changes in price – elasticity – but it is unlikely that cheap carrots and apple slices will be good substitutes for chocolate chip cookies, even if they are more expensive. In a market economy we have the freedom to make bad choices. To what degree should governments restrict some choices (purchasing decisions) to improve health and well-being? Should we regulate or even ban certain food advertisements? Mandate how restaurants should serve food? Ban snack food from school vending machines?

*Climate change*

The largest and most significant negative externality is climate change. Individual actions that lead to carbon emissions can generate profound changes in the earth's temperature with potentially harmful global effects. How can our Ecological footprint – the resources, energy and space needed to provide our products and absorb our waste – be reduced?

Ecological economists present a stark departure from standard economics in the challenge they present to the fundamental goal of economic growth (expansion of our economy and increases in national income). They want to replace the objective of promoting economic growth with achieving 'optimal scale' (interpreted as an output level that is environmentally sustainable). They go further and advocate zero growth of the economy ('a steady state economy') as they believe we have exceeded the capacity of the ecosystem to support the insatiable demands we make on natural resources – energy, water, raw materials, land and waste products. We are close to resource exhaustion and the waste absorption capacity of the planet, they claim.

Followers of ecological economists like Herman Daly lament that we have degraded the world's fisheries and soils, poisoned its waters, and changed its climate, while driving countless other species to extinction. They note that we face rising temperatures and adverse weather conditions, with one quarter of the fisheries overfished, along with extensive deforestation, oil and fresh water shortages, while chemicals damage the ozone layer causing higher skin cancer rates.

These concerns raise broader issues of whether growth is sustainable and does it improve the quality of our lives? Is limitless growth an illusion and does endless material gain promote well-being? What are our objectives, as individuals or as a society – material wealth, leisure, solid relationships, good health and education opportunities, secure retirement, happiness, a sustainable environment, and a fair wealth distribution? Instead of using Gross Domestic Product, Bhutan uses an 'Gross National Happiness' index incorporating psychological well-being, physical health, work-life balance, community vitality and social connection, education, cultural preservation and diversity, sustainability, good governance and material well-being. Should other countries do the same? These issues raise fundamental issues relating to how governments intervene, regulate and shape societies.

# ENDNOTES

1. The convexity of the IDCs can be explained in terms of declining absolute value of the slope of the IDCs as quantities of X rise and quantities of Y fall.
2. The relationship between average and marginal can be shown using simple algebra.
   The slope of AP = dAP/dL = d(Q/L) / dL = (dQ/dL × L – Q/L ) / $L^2$ = 1/L (MP – AP).
   1/L is always positive. The slope of AP > 0 when MP > AP, that is AP is rising. The slope of AP < 0 when MP < AP, that is AP is falling. The slope of AP = 0 when MP = AP, that is AP is at maximum. This is a general relationship which applies to all of the variables.
3. It can be shown that the slope of an isoquant is the ratio of marginal products of two inputs.
   For an isoquant Q = f (k,L) = C where C is a constant. $dQ = \frac{\partial Q}{\partial K} dK + \frac{\partial Q}{\partial L} dL = 0$.
   Solving that equation for – dy /dx = $MP_L$ / $MP_K$.
   It can be easily shown that the slope of an isocost is w / r.
4. Remember that at the point of tangency two curves have equal slopes. If we assume that LRAC is U shape then all of the ATC curves to the left of the minimum of LRAC must be tangent before their minimum and all of those tangents to the right of the minimum of LRAC must be above their respective minimums.
5. The AR curve and the demand curve are inverse function. The demand curve is quantity as a function of price whereas the AR is price as a function of quantity.
6. In Figure 7.1 the VMP curve is the demand for variable input when there is only one input. When a firm uses several variable inputs, at the lower wage, the quantity of one input and its complements rises causing shift of the VMP curve to the right. Overall, at the lower wage employment of one input with several inputs is larger than if there was only one input. However, in both cases the demand for labour is downward sloping.
7. Needless to say the technical discussion of factor pricing covered in the first half of this chapter may seem a little dry and arid but the consequences for income inequality are highly contestable and controversial.