UNIVERSITY OF LONDON | INTERNATIONAL PROGRAMMES

Industrial economics

G. Symeonidis

EC3**099**, 2790**099**

2011

Undergraduate study in
**Economics, Management,
Finance and the Social Sciences**

LSE THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE ■

Contents

Notes

Introduction

This subject guide provides an introduction to current theory and empirical work in industrial economics. It starts by examining the internal structure of firms. It then moves on to the analysis of various aspects of strategic interaction between firms and the determinants of industrial structure. Finally, it discusses the role of policy in the context of competition and industrial policies and regulation. The emphasis throughout will be on understanding how the theoretical tools can be used to analyse real-world issues. The theory will be confronted against empirical evidence, and its implications for public policy and business strategy will be discussed.

Aims and objectives

This course aims to:

- provide you with the analytical skills required for understanding problems in industrial economics, including applications of game theory

- examine the key questions on the internal organisation of firms

- analyse various aspects of strategic interaction between firms and the determinants of industrial structure

- provide you with the ability to apply economic models of firm behaviour to analyse questions in business strategy, competition policy and regulation.

Learning outcomes

At the end of the course and having completed the Essential reading and Activities you should be able to:

- describe and explain the determinants of the size and structure of firms and the implications of the separation of ownership and control

- describe and explain the pricing behaviour by firms with market power and its welfare implications

- apply analytical models of firm behaviour and strategic interaction to evaluate various business practices, including tacit collusion, entry deterrence, product differentiation, price discrimination and vertical restraints

- recognise and explain the basic determinants of market structure and the key issues in competition policy and regulation.

Syllabus

Theory of the firm

Size and structure of firms: the technological view of the firm; the transaction costs–property rights approach; investment specificity, incomplete contracts and vertical integration; empirical evidence.

Separation of ownership and control: separation of ownership and control; managerial incentives; the limits to managerial discretion; foundations of the profit-maximisation hypothesis.

Firm conduct and market structure

Short-run price competition: the Bertrand model; Bertrand competition with capacity constraints; the Cournot model.

Dynamic price competition: repeated interaction; collusion and cartel stability; theories of price wars; empirical analysis of market power and collusive behaviour.

Entry deterrence and entry accommodation: first-mover advantages and the value of irreversible decisions; strategies to deter entry; strategic substitutability vs. complementarity; a taxonomy of business strategies.

Product differentiation and non-price competition: horizontal product differentiation; brand proliferation and entry deterrence; vertical product differentiation.

Price discrimination: first-degree, second-degree and third-degree price discrimination; non-linear pricing; tie-in sales.

Vertical restraints: efficiency explanations for vertical restraints; vertical and horizontal externalities; vertical restraints as instruments that restrict competition; empirical evidence.

The determinants of market structure: theory of market structure in exogenous and endogenous sunk cost industries; technology and market structure; empirical evidence.

Competition policy and regulation

Competition and industrial policy: competition policy in the EU, the USA and Japan; current issues in competition policy; industrial policy towards R&D.

Regulation: regulation of firms with market power under symmetric information; regulation under asymmetric information; liberalisation and regulation; empirical evidence.

Prerequisites for this subject

If you are taking this course as part of a BSc degree you must have taken either **28 Managerial economics** or **66 Microeconomics** first.

Knowledge of microeconomic analysis at an intermediate level is necessary for students taking **99 Industrial economics**. This subject guide assumes that you are fully familiar with the theory of costs, the analysis of alternative market structures such as perfect competition, monopoly and oligopoly at an intermediate level, and concepts from consumer theory and welfare economics.

99 Industrial economics makes considerable use of game theory. For the purposes of this subject some knowledge only of elementary game theory is required. The emphasis is on using the game-theoretic techniques in applications, so you do not need to worry about abstract definitions. You must be familiar with the concepts of Nash equilibrium in static games and subgame-perfect equilibrium in dynamic games. However, you do not need to have any knowledge of games with incomplete information for this course.

Finally, the algebra required is simple calculus. Most of the mathematical problems you will be faced with in this course are simple maximisation problems. Regarding statistics, only some knowledge of elementary probability theory is required. Some familiarity with basic econometric techniques will help you better understand and appreciate some of the empirical readings, although it is not essential for the examination.

The structure of this subject guide

This subject guide is divided into three parts. Part 1 analyses topics in the theory of the firm. Chapter 1 looks at the theory and evidence on the factors determining the size and structure of firms. It reviews the technological view of the firm, and then focuses on the transaction costs–property rights approach. The links between investment specificity, contracts and vertical integration are discussed. The chapter ends with a review of empirical evidence.

Chapter 2 is concerned with the implications of the separation of ownership and control in modern large firms. A number of issues are examined, including managerial incentives, the limits to managerial discretion, and the foundations of the profit-maximisation hypothesis.

The next part examines various aspects of oligopolistic interaction. **Chapter 3** is about short-run competition between firms. Alternative models are discussed and compared, including the Bertrand model, the model of price competition with capacity constraints, and the Cournot model.

Chapter 4 looks at dynamic price competition. Topics analysed include repeated interaction, collusion and cartel stability, and theories of price wars. The chapter concludes with a discussion of empirical analyses of market power and collusive behaviour.

Chapter 5 focuses on situations where firms can act strategically to influence the decisions of other firms. It examines first-mover advantages and the value of irreversible decisions, strategies to deter entry, and concludes with a useful framework for classifying business strategies.

Chapter 6 addresses the specific issues associated with product differentiation and non-price competition. It analyses both horizontal and vertical product differentiation, and shows, in the context of a particular case study, how this analysis can help understand brand proliferation as a strategy to deter entry.

Chapter 7 reviews the various types of price discrimination by firms and discusses applications, such as the use of tie-in sales.

Chapter 8 focuses on vertical relationships between firms and reviews several reasons why firms may want to use vertical restraints. Explanations of vertical restraints that emphasise efficiency gains and others that emphasise welfare losses through the restriction of competition are analysed. Empirical evidence on vertical restraints is discussed.

This part concludes with an analysis of the determinants of market structure in **Chapter 9**. Specific topics include the links between competition and market structure, and technology and market structure. The theory is tested against empirical evidence.

The final part of the guide focuses on public policy towards industry. **Chapter 10** reviews competition policy in the EU, the USA and Japan, and examines the main economic issues in the design and implementation of competition policy. Industrial policy towards R&D is also discussed.

Chapter 11 examines topics in the regulation of firms with market power, both under symmetric information and under asymmetric information. It also examines competition in regulated industries, and provides some empirical evidence on these issues.

How to use this subject guide

This subject guide has been written in such a way that you can obtain a basic understanding of the topics in the syllabus from the guide, before going on to read the various texts to broaden and deepen this understanding. This is necessary because there is no single text that covers all the topics and also because some of the best texts in industrial economics are at a level rather more advanced than is required for this course. The subject guide was therefore also designed to give you a clear indication of the level of analysis that will be expected of you in the examination.

The guide is, however, not a substitute for the careful study of the readings listed at the beginning of each chapter. Its purpose is to help you organise your study and give you a starting point for each topic in the syllabus. While there is no single best way to organise your study for this course, it may be useful, for each topic in the syllabus, to start with the relevant chapter of the guide, then do the recommended reading for that particular topic, then come back to the guide and attempt the questions at the end of the chapter.

Each chapter of the guide contains 'learning outcomes' to enable you to check your progress. Each chapter also contains Activities and sample examination questions. There are two types of activity. Some are self-test questions designed to test your understanding of the material contained in the guide and/or the recommended reading. You are strongly encouraged to attempt these questions as well as the sample examination questions. Simply reading the guide and the recommended texts is not sufficient for maximising your benefit from the subject or for preparing for the examination.

Other Activities ask you to study additional material (including internet resources) and are mainly designed to deepen your understanding of the links between economic theory and empirical evidence: case studies, results from laboratory experiments and mini-surveys of the evidence on particular issues. These activities are intended as optional. They are there to give you the opportunity to find out more about topics of interest to you.

Essential reading

Books

There is no single text that covers all the topics in the syllabus. The list below contains several references. An excellent reference for many of the topics covered is:

> Tirole, J. *The Theory of Industrial Organization.* (Cambridge, MA: MIT Press, 1988) [ISBN 9780262200714].

This book mainly focuses on theory. Moreover, certain parts of it contain material which is too advanced for this course. As a general rule, you do not have to worry about any material contained in the appendices. Finally, not all chapters of the book are relevant for the syllabus of this course. However, Tirole's text does an excellent job in combining simple formal economic analysis with a rich informal discussion of many important issues in industrial economics. Several of the chapters in the subject guide use or build upon examples taken from Tirole's book.

Another very good reference which covers most topics in the syllabus at a level less advanced than Tirole's text is:

Church, J.R. and R. Ware *Industrial Organization: A Strategic Approach.* (Irwin McGraw-Hill, 2000) [ISBN 9780256205718].

This book emphasises strategic behaviour and covers a wide range of topics in great detail (once again, not all of these topics are included in the syllabus of this course). It also contains a large number of case studies to motivate or illustrate the economic analysis. At the time of writing of this guide it was available for free download from the internet: http://homepages.ucalgary.ca/~jrchurch/page4/page5/files/PostedIOSA.pdf

Two other books are useful for particular parts of the syllabus. For Chapter 9, you may refer to:

Sutton, J. *Sunk Costs and Market Structure*. (Cambridge, MA: MIT Press, 1991) [ISBN 9780262193054].

And for Chapter 11, you may read certain parts of:

Armstrong, M., S. Cowan and J. Vickers *Regulatory Reform*. (Cambridge, MA: MIT Press, 1994) [ISBN 9780262011433)].

Journals

A number of journal articles or book chapters are included under the category of 'Essential reading' because they are discussed in some detail in the guide. They are the following:

Porter, R.H. 'A Study of Cartel Stability: The Joint Executive Committee, 1880–1886', *Bell Journal of Economics* (1983) 14(2), pp.301–314.

Rey, P. and J. Stiglitz 'The Role of Exclusive Territories in Producers' Competition', *Rand Journal of Economics* (1995) 26(3), pp.431–451.

Schmalensee, R. 'Entry Deterrence in the Ready-to-eat Breakfast Cereal Industry', *Bell Journal of Economics* (1978) 9(2), pp.305–327.

Sutton, J. 'Market Structure: Theory and Evidence', in Armstrong, M. and R. Porter (eds) *Handbook of Industrial Organization*, volume 3. (Amsterdam: North-Holland, 2007) [ISBN 9780444824356]. Working paper version available at: http://personal.lse.ac.uk/sutton/market_structure_theory_evidence.pdf

Further reading

Please note that as long as you read the Essential reading you are then free to read around the subject area in any text, paper or online resource. You will need to support your learning by reading as widely as possible and by thinking about how these principles apply in the real world. To help you read extensively, you have free access to the virtual learning environment (VLE) and University of London Online Library (see below).

There are several other textbook references for this subject. Some of the most useful are:

Cabral, L. *Introduction to Industrial Organization.* (Cambridge, MA: MIT Press, 2000) [ISBN 9780262032865].

Carlton, D.W. and J.M. Perloff *Modern Industrial Organization.* (Pearson Addison Wesley, 2005) fourth edition [ISBN 9780321223418].

Martin, S. *Industrial Economics* (Englewood Cliffs, NJ: Prentice Hall, 1994) second edition [ISBN 9780023767869]

Shy, O. *Industrial Organization.* (Cambridge, MA: MIT Press, 1995) [ISBN 9780262691796].

The first three of these can be used to complement Tirole's more formal analysis with a more descriptive, yet rigorous, treatment of the various

topics. Cabral's text is relatively concise, but still covers a lot of ground, and contains a good discussion of the modern literature on market structure. A new textbook by Martin, *Industrial Organization in Context*, is due to be published in 2009. It will cover both economic theory and recent developments in competition policy in the US and the EU. Shy's text is analytically more advanced than the other three, with an emphasis on simple formal economic models.

A good collection of case studies in antitrust (competition) policy is:

Kwoka, J.E. and L.J. White (eds) *The Antitrust Revolution: Economics, Competition, and Policy*. (New York: Oxford University Press, 2004) fourth edition [ISBN 9780195161182].

There is also a companion website where you can download all the case studies published in previous editions of this book: www3.oup-usa.org/sc/0195161181/

Finally, you may wish to also refer to:

Scherer, F.M. and D. Ross *Industrial Market Structure and Economic Performance.* (Boston: Houghton Mifflin, 1990) third edition [ISBN 978-0395357149].

This book covers empirical, institutional and public policy issues quite extensively (but mostly with reference to the United States) and is a good complement to any text that focuses primarily on theory.

All the above books are useful for the course as a whole or as primary references for particular topics. Additional references for particular topics of the syllabus will be given at the beginning of each chapter.

Most current research in industrial economics is published in academic journals. You have free access to a number of these via the University of London Online Library. In addition to the numerous general Economics journals which regularly publish articles in industrial economics, there are a number of specialist journals, including:

- *Journal of Industrial Economics*
- *Journal of Economics and Management Strategy*
- *International Journal of Industrial Organization*
- *Rand Journal of Economics.*

Information on competition issues – including news, reports and case studies – is posted on the websites of competition authorities. For instance:

- European Union: http://ec.europa.eu/comm/competition/index_en.html
- United Kingdom: www.competition-commission.org.uk/ (Competition Commission) and www.oft.gov.uk/ (Office of Fair Trading)
- United States: www.ftc.gov/bc/index.shtml (Federal Trade Commission) and www.usdoj.gov/atr/ (Antitrust Division, Department of Justice).

Online study resources

In addition to the subject guide and the Essential reading, it is crucial that you take advantage of the study resources that are available online for this course, including the VLE and the Online Library.

You can access the VLE, the Online Library and your University of London email account via the Student Portal at:

http://my.londoninternational.ac.uk

You should receive your login details in your study pack. If you have not, or you have forgotten your login details, please email uolia.support@ london.ac.uk quoting your student number.

The VLE

The VLE, which complements this subject guide, has been designed to enhance your learning experience, providing additional support and a sense of community. It forms an important part of your study experience with the University of London and you should access it regularly.

The VLE provides a range of resources for EMFSS courses:

- Self-testing activities: Doing these allows you to test your own understanding of subject material.

- Electronic study materials: The printed materials that you receive from the University of London are available to download, including updated reading lists and references.

- Past examination papers and *Examiners' commentaries*: These provide advice on how each examination question might best be answered.

- A student discussion forum: This is an open space for you to discuss interests and experiences, seek support from your peers, work collaboratively to solve problems and discuss subject material.

- Videos: There are recorded academic introductions to the subject, interviews and debates and, for some courses, audio-visual tutorials and conclusions.

- Recorded lectures: For some courses, where appropriate, the sessions from previous years' Study Weekends have been recorded and made available.

- Study skills: Expert advice on preparing for examinations and developing your digital literacy skills.

- Feedback forms.

Some of these resources are available for certain courses only, but we are expanding our provision all the time and you should check the VLE regularly for updates.

Making use of the Online Library

The Online Library contains a huge array of journal articles and other resources to help you read widely and extensively.

To access the majority of resources via the Online Library you will either need to use your University of London Student Portal login details, or you will be required to register and use an Athens login: http://tinyurl.com/ollathens

The easiest way to locate relevant content and journal articles in the Online Library is to use the **Summon** search engine.

If you are having trouble finding an article listed in a reading list, try removing any punctuation from the title, such as single quotation marks, question marks and colons.

For further advice, please see the online help pages: www.external.shl.lon.ac.uk/summon/about.php

Examination advice

Important: the information and advice given here are based on the examination structure used at the time this guide was written. Please note that subject guides may be used for several years. Because of this we strongly advise you to always check both the current *Regulations* for relevant information about the examination, and the VLE where you should be advised of any forthcoming changes. You should also carefully check the rubric/instructions on the paper you actually sit and follow those instructions.

Remember, it is important to check the VLE for:

- up to date information on examination and assessment arrangements for this course
- where available, past examination papers and *Examiners' commentaries* for the course which give advice on how each question might best be answered.

This subject is assessed by a three-hour examination. You will be required to answer four questions out of eight. A sample examination paper is included at the end of this guide. As you will see from that paper, some questions will be problem-type questions, while others will be essay-type questions. Problem-type questions are quite specific as to what you are required to do. Essay-type questions are sometimes less specific, but a good answer to an essay-type question must include some rigorous economic analysis, usually with reference to some economic model or models.

You can find guidance on examination technique in the annual *Examiners' commentaries* for the course which are available on the University of London International Programmes website, and also in your academic and study skills handbook, *Strategies for success*.

Chapter 1: Size and structure of firms

Learning outcomes

By the end of this chapter, and having completed the Essential readings and Activities, you should be able to:

- describe and evaluate two different approaches to explaining the size and structure of firms
- explain the notions of 'transaction costs', 'investment specificity', 'opportunistic behaviour', 'incomplete contracts' and 'residual rights of control' and their relevance for the theory of the firm
- analyse two different types of inefficiency that can arise in the context of long-run relationships between firms
- explain the effect of investment specificity on the decision of firms to enter into contractual relationships or to integrate.

Essential reading

Church, J.R. and R. Ware *Industrial Organization: A Strategic Approach.* Chapter 3.

Tirole, J. *The Theory of Industrial Organization.* Introductory chapter.

Further reading

Books

Hart, O. *Firms, Contracts, and Financial Structure.* (Oxford: Oxford University Press, 1995). Chapters 1–4.

Perry, M.K. 'Vertical Integration: Determinants and Effects', in Schmalensee, R. and R. Willig (eds) *Handbook of Industrial Organization*, volume 1. (Amsterdam: North-Holland, 1989).

Journals

Hart, O. 'An Economist's Perspective on the Theory of the Firm', *Columbia Law Review* (1989) 89(7), pp.1757–1774.

Hart, O. and J. Moore 'Property Rights and the Nature of the Firm', *Journal of Political Economy* (1990) 98(6), pp.1119–1158.

Joskow, P.L. 'Contract Duration and Relationship-Specific Investments: Empirical Evidence from Coal Markets', *American Economic Review* (1987) 77(1), pp.168–185.

Klein, B., Crawford, R.G. and Alchian, A.A. 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process', *Journal of Law and Economics* (1978) 21(2), pp.297–326.

Lafontaine and Slade 'Vertical Integration and Firm Boundaries: The Evidence', *Journal of Economic Literature* (2007) 45(3), pp.629–685.

Monteverde, K. and D.J. Teece 'Supplier Switching Costs and Vertical Integration in the Automobile Industry', *Bell Journal of Economics* (1982) 13(1), pp.206–213.

Introduction

What explains the size and structure of firms? In fact, you may also ask: why do agents group together to form firms? This chapter aims to provide some answers to these questions and also to examine some more specific issues, namely why some transactions take place within firms while others are conducted through external contractual relationships; what determines, in the case of contracts between firms, the types of contracts used; and finally, what are implications of alternative ownership structures are for efficiency.

This chapter focuses on **efficiency explanations** for the size and structure of firms. It is important to understand what 'efficiency' means in this context. Efficiency motives are those associated with minimising costs or maximising producer surplus in a way that may also be socially beneficial, that is increase total social welfare. The efficiency of a certain organisational form refers then here primarily to the firm or firms involved, not necessarily to society as a whole.

You should bear in mind that there are also market power explanations for the size and structure of firms. Unlike efficiency motives, market power motives induce behaviour which, although profitable for the firm or firms involved, is definitely detrimental to social welfare. For instance, two firms producing the same product may merge not to reduce costs but simply to enhance their ability to exercise market power. Much of the second part of this subject guide is concerned with the behaviour of firms with market power. So to make an overall assessment of the factors that determine the size and structure of firms you should first work through most of the guide.

There are two broad classes of efficiency explanations: the technological view of the firm and the transaction costs–property rights approach.

Technological factors

According to the technological view, optimal firm size and diversification depend on the degree of economies of scale and scope.[1] For instance, a single-product firm may have an average cost curve such as the one depicted in Figure 1.1. To minimise average cost, the firm will in this case operate at a size between q_1 and q_2.
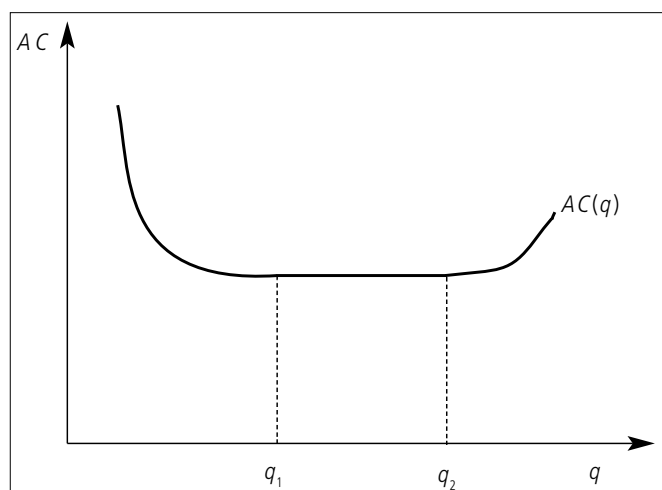
[1] See Tirole (1988), pp.18–21 for details. On the concepts of scale and scope economies, you can also read Church and Ware (2000), Chapter 3.

Figure 1.1

While technological constraints are important, they are not the whole story. In particular, there are two problems with the technological view of the firm:

- It may explain the joint use of facilities, but not joint ownership. Why can't agents write contracts to exploit economies of scale and scope without joint ownership?

- It is not clear why the AC curve rises at high output. If producing quantity $q_A + q_B$ were to cost more than producing q_A and q_B separately, why can't there be a single firm that consists of two independent divisions producing q_A and q_B respectively?

The transaction costs–property rights approach

Transaction costs, incomplete contracts and integration

The starting point for this approach is the idea that the choice between organising activity internally and using the market (or contracts) is determined by a comparison of the costs and benefits of these two modes of organisation. Williamson has identified some economic factors that matter for this choice. There are three elements in his approach:

1. investment specificity

2. opportunistic behaviour

3. bounded rationality.

Many long-run relationships between economic agents involve relation-specific investments, in other words investments that pay off a maximum return only if the particular relationship continues for some time. Examples of specificity are site specificity (e.g. a firm builds a plant next to another firm's works to save transport costs); physical asset specificity (e.g. a firm designs equipment the characteristics of which are specific to a particular order); and human capital specificity (e.g. an employee invests in acquiring skills which are specific to a certain job). In all these cases, the first best use value of an investment is higher than its value in any alternative use.

Now *ex ante* (i.e. before any investments are made) there is a competitive situation. For instance, if the relationship is between a buyer and a supplier of a certain product, there will be many suppliers and buyers and they can select each other out of the pool of competitive suppliers and buyers. But *ex post* (i.e. after investments have been made) there is a bilateral monopoly, because if the parties trade with each other they can make gains which will not be made otherwise. This creates the possibility of a 'hold-up' or opportunistic behaviour. Each party wants to appropriate the common surplus *ex post*, so there will be bargaining. This can create several problems, in addition to any costs of haggling:

- The level of trade ex post may not be efficient if there is asymmetric information. This can occur irrespective of whether the relation involves ex ante investments or not.

- The level of investment ex ante will not be efficient, even under symmetric information. The reason is that once a party has sunk the cost of the investment, it has lost any extra bargaining power. So even if the efficient volume of trade occurs ex post, the division of the surplus will be such that the level of investment ex ante is not efficient.

These notions will be made more precise with the help of a formal model below. But before that, consider the following question. If the parties could write a contract *ex ante* specifying the terms of trade *ex post*, would they be able to achieve efficiency? The answer is that they might. So why can't they write such a contract? This is where the 'bounded rationality' idea comes in: a complete contract is impossible to write. There are several reasons for this, such as unforeseen contingencies, prohibitive costs of contracting over all contingencies (even if these are foreseeable), or prohibitive monitoring costs. As a result, contracts are necessarily incomplete. Some bargaining will have to take place *ex post*, and this may lead to inefficiencies.

We can now state an important theoretical prediction of the transaction costs approach: the more specific the investment, the larger the scope for efficiency losses due to opportunistic behaviour. Hence the more specific the investment, the higher the probability of integration (i.e. common ownership) as opposed to a contractual relationship.

A model

Let us formalise some of the above ideas using a simple model of a vertical relationship between a buyer and a seller.[2] There are two periods, t = 1 (*ex ante*) and t = 2 (*ex post*). We first want to focus on the issue of *ex post* efficiency, so assume for now that there is no investment *ex ante*. The two parties can, if they wish, trade one unit of an indivisible good in period 2. Let v denote the value of the good to the buyer (this can be the difference between the value to the buyer in this relationship and that in an alternative relationship), c the production cost, and p the price at which the parties trade. If there is trade, the buyer has a gain of $v - p$, while the seller has a gain of $p - c$. If there is no trade, both gain zero.

If v and c are known to both parties at the beginning of period 2, then the volume of trade is efficient, which is another way of saying that trade will occur if and only if there are gains from trade, in other words if and only if $v \geq c$. That is so because, if $v \geq c$, the parties will agree to trade at a price p such that $c \leq p \leq v$ rather than make zero surplus.[3] While if $v < c$, at least one party would be making negative surplus if trade were to occur, so this party will refuse to trade. More generally, it can be shown that under symmetric information we always obtain the *ex post* efficient outcome.

If, however, there is asymmetric information, the volume of trade may be inefficient. Suppose that both parties know c but only the buyer knows v. All the supplier knows is that v is distributed as a random variable with cumulative distribution function $F(v)$ and density function $f(v)$ on the interval $[\underline{v}, \overline{v}]$ (hence $F(\underline{v}) = 0$, $F(\overline{v}) = 1$).[4] Gains from trade exist with some probability between 0 and 1, that is $\underline{v} < c < \overline{v}$ (the problem would be trivial otherwise). To simplify the problem suppose further that the supplier has all the bargaining power in period 2, so he makes a take-it-or-leave-it offer to the buyer at price p.[5] The buyer will accept this offer if $v \geq p$, so trade will occur if and only if $v \geq p$, that is with probability:

$$prob(v \geq p) = \int_{p}^{\overline{v}} f(s)ds = 1 - F(p)$$

Recall that if trade does not occur the supplier ends up with zero. So the supplier's expected profit is given by $E(\Pi) = (p - c)[1 - F(p)]$. The supplier will choose p to maximise this, so:

$$\frac{\partial E(\Pi)}{\partial p} = 1 - F(p) - (p - c)f(p) = 0$$

From this equation it can be seen that in general the supplier chooses $p > c$, so there are circumstances where trade does not occur even though there are gains from trade. In particular, this is the case when $p > v \geq c$. In this case the buyer rejects the supplier's offer since he would make a loss by accepting, so trade does not occur even though $v \geq c$. For efficiency, on the other hand, we would require $p = c$, so that trade occurs if and only if $v \geq c$. More generally, it can be shown that when both value and cost are private information and gains from trade are not certain, the volume of trade is not efficient.

Activity

Show that, if the parties could sign a contract in period 1 in this simple model, they could devise a contract such that the efficient volume of trade is obtained.

Answer

The contract should simply give the power to choose the price p in period 2 to the buyer (i.e. the informed party). The buyer would then set $p = c$ in order to appropriate all the surplus.[6] So trade would occur if and only if $v \geq c$, which is what we require for efficiency. The fact that the buyer would appropriate all the surplus is irrelevant as far as efficiency is concerned, because all that efficiency requires is the maximisation of the 'pie'. In any case, the parties could also specify in the contract a lump sum payment from the buyer to the seller to create any division of the surplus: the outcome would depend on the relative bargaining power of the parties in period 1, when the contract is signed. Note, however, that, unlike our simple model, a complete contract cannot be written in many practical situations. In those cases, asymmetric information will lead to inefficient outcomes.

So far there was no *ex ante* relation-specific investment in the model. Now let us assume that one of the parties, say the supplier, can invest **in period 1**, say in cost reduction. In particular, let c be a function of investment I: $c(I)$, with $c'(I) < 0$, $c''(I) > 0$. Assume $v \geq c(0)$ to ensure positive optimal investment. We want to focus on how *ex post* bargaining affects the volume of investment undertaken, so we will further assume that there is symmetric information and hence the volume of trade *ex post* is efficient. In other words we abstract from additional complications created by asymmetric information. So both v and c are commonly known at the beginning of period 2 and trade occurs if and only if $v \geq c$. Finally assume for simplicity that the two parties have no 'outside option' *ex post*, that is to say their only chance to realise a positive surplus is to trade between themselves (there are no other buyers or sellers in period 2).

In period 2 the trading price p will be determined through bargaining. If the two parties have equal bargaining power, the *ex post* surplus will be split equally between them, namely:

$$v - p = p - c(I) \Leftrightarrow p = \frac{v + c(I)}{2}$$

Recall that v – p is the buyer's *ex post* surplus (i.e. his gain over and above his second-best alternative, which is in this case zero), while $p - c(I)$ is the seller's *ex post* surplus. Note that the cost of investment I is not relevant as far as the division of the *ex post* surplus is concerned because this investment has already been sunk when the two parties bargain.

Now at date 1 the supplier chooses how much to invest. When making this decision he anticipates what will happen in period 2, that is, he anticipates that $p = [v + c(I)]/2$. He chooses I to maximise his net profit, which is equal to the *ex post* surplus minus the cost of investment:

[6] *Why is the supplier prepared to trade if $p = c$? Recall that to avoid some purely technical problems we have assumed that when a party is indifferent between trading and not trading, it chooses to trade. Alternatively, you can imagine that the buyer offers $p = c + \varepsilon$, where e is positive and very small.*

$$p - c(I) - I = \frac{v + c(I)}{2} - c(I) - I$$

The first-order condition is $-c'(I) = 2$, and this implicitly defines the privately optimal level of investment I_p. Is this level of investment efficient? The efficient level of investment is the value of I that maximises the joint net profit $v - c(I) - I$. In other words it is the value of I that maximises the 'pie' net of the cost of investment. You can also think of it as the value of I that would be chosen if the supplier and the buyer merged into a single entity. The first-order condition is $-c'(I) = 1$ and this defines the efficient level of investment I^*. Since $c''(I) > 0$ (i.e. the cost function is strictly convex) we have $I_p < I^*$ (see Figure 1.2). The supplier invests less than what is required for efficiency.
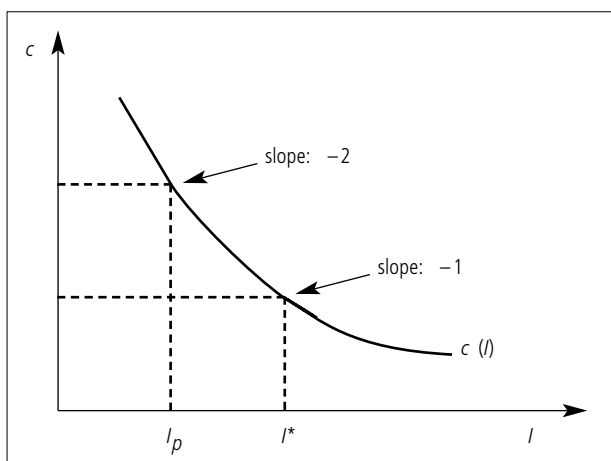
Figure 1.2

The intuition is simple. Since the *ex post* surplus is divided between the two parties, the investing party does not capture all the cost savings from its investment. This 'distortion' of incentives leads to underinvestment. The model can also be refined to analyse the effect of the degree of investment specificity on the level of *ex ante* investment. It turns out that the level of investment is lower the higher the degree of specificity. Thus the higher the degree of specificity, the bigger is the incentive for the firms to merge if a contractual solution is not feasible. (Actually in our example a contractual solution is feasible, but in more general settings it would not be.)

The above analysis has left some questions unanswered, however. Exactly why does integration solve or reduce the hold-up problem (i.e. exactly what changes when two firms merge?). And why then don't firms always merge (i.e. why are there limits to integration?).

Property rights

Hart and others have pointed out that, given that contracts are incomplete, one thing that greatly matters in a relationship is which party has the right to make decisions in the case of unspecified contingencies. Obviously, it is the owner of the physical asset(s) who has this right, the 'residual right of control'. According to this view, a firm is seen as a collection of physical assets that belong to it: machines, inventories, buildings, client lists, patents, cash, etc. – excluding human capital. 'Ownership' is defined as the right to specify all usages of these assets in any way not inconsistent with a previous contract, custom or law. Note that the possession of residual rights of control does not rule out *ex post* renegotiation. What it does is determine the 'status quo point' in the bargaining process, in other words

it puts the party that has these rights in a better bargaining position. In this way it affects the division of the surplus *ex post* and therefore also influences the level of investment *ex ante*.[7]

Several conclusions have emerged from the property rights approach. Consider the case of two owner-managed firms that enter into a long-run relationship and must both make a relation-specific investment. Then:

- Integration reduces opportunistic behaviour because if, say, firm A acquires firm B, then the manager of firm B loses control of the physical assets of firm B, so he has much less bargaining power.

- A party is more likely to own an asset if it has a large investment to make. In other words, efficiency requires that the residual rights of control rest with the party whose ex ante investment has the larger effect on the profits of both parties.

- Efficiency requires that highly complementary assets are under common ownership. On the other hand, independent assets should be separately owned: there are limits to integration. Why? If firm C acquires firm D, say, then the manager of firm D will have much lower incentives to undertake investments since the payoff from these will be partly appropriated by the owner of firm C. So if assets are independent, the costs of integration (in terms of underinvestment) will be higher than any potential benefits.

You should also bear in mind some qualifications. First, the above conclusions apply more directly to the case of owner-managed firms. Things are more complicated when there is separation of ownership from control and delegation of authority within firms, although the general framework should still be valid. Second, another way of solving the problem of opportunism may be for firms to try to build a reputation for non-opportunistic behaviour when they interact repeatedly with each other. However, this would not necessarily work because the payoff from behaving opportunistically might be bigger than the payoff from adhering to non-opportunistic behaviour.[8]

Lowering transaction costs and opportunistic behaviour is not the only reason for vertical integration (i.e. integration between a supplier and a buyer). As we will see in later chapters of this guide, a firm may also vertically integrate to eliminate negative externalities that arise in buyer–seller relationships in the absence of relation-specific investments, or to price-discriminate, or to increase its market power by hindering the access of rival non-integrated firms to outlets or sources of supply.

Empirical evidence

Much of the empirical work on the determinants of firm size and structure that has followed the transaction costs approach has focused on the role of investment specificity for vertical integration.[9] Two examples are discussed below. Lafontaine and Slade (2007) provide a survey of the empirical literature on the boundaries of the firm.

Klein, Crawford and Alchian (1978) describe the story of the 1926 merger between General Motors (GM) and Fisher Body. In 1919 GM, a US car manufacturer, entered a 10-year contract with Fisher Body for the supply of car bodies. To minimise the scope for opportunistic behaviour the contract specified that GM should buy all their closed car bodies from Fisher and also specified the trading price with the additional provision that this price could not be greater than the average market price of similar bodies produced by firms other than Fisher.

[7] *Hart and Moore (1990), in their Introduction, discuss a similar example with one asset and three agents which illustrates many of the main ideas and results in the property rights approach. Tirole (1988), pp.29–34, and Hart (1995), Chapters 2–4, provide a more formal treatment as well as rich informal discussion.*

[8] *This point will become clearer when you have studied the material on repeated interaction in Chapter 4 of this guide.*

[9] *Other studies have focused on the related issue of the role of investment specificity for the type and duration of contracts signed between firms. Joskow (1987) is a good example.*

However, demand conditions changed dramatically after 1919: there was a large increase in demand for cars, especially cars with closed bodies – the type manufactured by Fisher. GM thought that Fisher's cost had gone down because of scale economies in the production of bodies and were unhappy with the price they had to pay for Fisher bodies. Also, Fisher refused to locate their plants close to GM plants – a move which GM thought would increase production efficiency but which would diminish the bargaining power of Fisher. These tensions ended in 1926, when GM acquired Fisher.[10]

[10] This is the prevailing view on the GM–Fisher merger. See, however, Activity 3 below.

Monteverde and Teece (1982) have examined why firms in the automobile industry produce some components in-house while they buy others from independent suppliers. One of their main hypotheses was that car manufacturers will vertically integrate when the production process for components generates transaction-specific know-how. That is so because it is then more difficult for them to switch to other suppliers, so there is more scope for opportunistic behaviour by suppliers.

Monteverde and Teece tested this hypothesis econometrically using data on 127 different components used by two big US car manufacturers. The dependent variable in their regressions was a binary variable for in-house production versus production by an external supplier. Their independent variables included the cost of developing a component (a proxy for the know-how generated in the production of a component), a dummy variable for firm-specific components versus generic components, a firm dummy to control for company effects, and other variables. Their results confirmed the predictions of the transaction costs approach:

- The higher the development cost of a component, the more likely that production was in-house.

- Firm-specific components were more likely to be produced in-house than generic components.

Activities

1. Consider the model of a vertical relationship between a buyer and a seller analysed above. We have seen that when the supplier can invest in period 1 to reduce the production cost c, he chooses a level of investment which is not efficient. Could a contract between the parties restore the efficient outcome? When should this contract be signed and what should it specify? Assume that a contract which directly specifies the level of investment that the supplier is to undertake is not feasible because investment levels, although observable by the parties, are not verifiable in a Court.

2. During the 1980s and 1990s, there was a trend towards de-integration across many industries as well as a trend towards more flexible technologies. Could the two be related and in what way?

3. The Klein, Crawford and Alchian (1978) interpretation of the 1926 merger between General Motors and Fisher Body has been criticised by some economists, including the Nobel prize winner Ronald Coase, who have argued that the reason for the merger had nothing to do with opportunistic behaviour and hold-up problems. This debate is not just about a particular event in economic history. It is about one of the most frequently cited examples of market failure. If the critics are right, then market failure in vertical relations between firms may be less prevalent than the theory would lead us to believe.

   What do you think? Make up your own mind after reading different views on the GM–Fisher Body case. A good collection of articles – representing the different views and including contributions by some of the protagonists of the debate – has been published in the April 2000 issue of the *Journal of Law and Economics*. You can also search the internet for more.

A reminder of your learning outcomes

By the end of this chapter, and having completed the Essential readings and Activities, you should be able to:

- describe and compare two different approaches to explaining the size and structure of firms

- explain the notions of 'transaction costs', 'investment specificity', 'opportunistic behaviour', 'incomplete contracts' and 'residual rights of control' and their relevance for the theory of the firm

- analyse two different types of inefficiency that can arise in the context of long-run relationships between firms

- explain the effect of investment specificity on the decision of firms to enter into contractual relationships or to integrate.

Sample examination questions

1. Consider the following model of a vertical relationship between a buyer and a seller. There are two periods and the two parties can, if they wish, trade one unit of an indivisible good in period 2. Let v denote the value of the good to the buyer, c the production cost, and p the trading price. Assume that $c < 1/2$. Both c and v are commonly known at the beginning of period 2. The seller can invest in period 1 to increase the value of the good to the buyer (for instance, he can spend on R&D to increase the quality of the product). In particular, $v(I) = 3I - I2/2$. The level of investment I cannot be specified in a contract because it is not verifiable and therefore such a contract would not be enforceable in Court.

   a. What is the efficient level of investment?

   b. In the absence of any contract, what is the level of investment chosen by the seller if the *ex post* surplus is to be divided equally between the two parties? Explain why this level is not efficient.

   c. Suppose that the parties sign a contract which gives to the **seller** the right to choose the trading price in period 2 (i.e. after the investment has been made). What will be the level of I chosen by the seller?

   d. Now suppose that the parties sign a contract which gives to the **buyer** the right to choose the trading price in period 2. What will be the level of I chosen by the seller? What is your conclusion about who should have the power to decide the price in period 2? Explain the intuition for your results.

2. Analyse how investment specificity affects the ex ante incentives for investment when there is ex post bargaining over the surplus. Then explain how investment specificity and the incompleteness of contracts may affect the decision of a firm to vertically integrate and discuss briefly any relevant empirical evidence.

Notes

Chapter 2: Separation of ownership and control

Learning outcomes

By the end of this chapter and having completed the Essential reading and Activities, you should be able to:

- explain the implications of the separation of ownership and control in modern large companies
- analyse optimal incentive mechanisms offered by the owners of a firm to the firm's manager
- describe different mechanisms that may restrict managerial discretion and discuss their limitations
- assess the validity of the profit-maximisation hypothesis.

Essential reading

Church, J.R. and R. Ware *Industrial Organization: A Strategic Approach*. Chapter 3.

Tirole, J. *The Theory of Industrial Organization.* Introductory chapter.

Further reading

Books

Holmstrom, B.R. and J. Tirole 'The Theory of the Firm', in Schmalensee, R. and R. Willig (eds) *Handbook of Industrial Organization,* volume 1. (Amsterdam: North–Holland, 1989).

Journals

Abowd, J.M. and D.S. Kaplan 'Executive Compensation: Six Questions that Need Answering', *Journal of Economic Perspectives* (1999) 13(4), pp.145–168.

Nickell, S. 'Competition and Corporate Performance', *Journal of Political Economy* (1996) 104(4), pp.724–746.

Nickell, S., D. Nicolitsas and N. Dryden 'What Makes Firms Perform Well?', *European Economic Review* (1996) 41, pp.783–796.

Symeonidis, G. 'The Effect of Competition on Wages and Productivity: Evidence from the UK', *Review of Economics and Statistics* (2008) 90, pp.134–146.

Introduction

A common assumption in most economic theory is that firms maximise (expected) profits. This is probably what the owners of a firm would like to do. However, in modern large companies, it is not the shareholders who run the firm, but the managers, who are likely to have other objectives than profit maximisation.

This separation of ownership and control gives rise to several important issues. First, given that the owners typically have less information than the managers and cannot perfectly monitor the behaviour of the latter, how can they design incentive schemes that induce the managers to behave as much as possible according to their (the owners') interests? Second, given that such contracts are generally not perfect solutions to the problem of

managerial discretion, what other mechanisms are there that may limit the ability of managers to pursue their own objectives rather than those of the owners? Third, is profit maximisation a reasonable description of firm behaviour?

Managerial incentives

An obvious way for the owners to restrict managerial discretion is to offer monetary or other incentives to managers. Some important insights on the use of incentives can be drawn from a simple model of a firm run by a single manager.[1] The profit of the firm can take one of two values, Π_1 and Π_2, with $\Pi_1 < \Pi_2$. The manager chooses between two levels of effort, high and low (for simplicity: zero). His utility is $U = u(w - \Phi)$ if he makes high effort and $U = u(w)$ if he makes zero effort, where u is an increasing and strictly concave function, w is the manager's wage, and $\Phi > 0$ is the monetary disutility of high effort. Note that the strict concavity of u implies that the manager is risk averse.[2] Whether the firm makes Π_1 or Π_2 depends on the manager's effort as well as on the firm's environment, which is uncertain. In particular, if the manager makes high effort, the profit is Π_2 with probability x and Π_1 with probability $1 - x$. If the manager makes no effort, the profit is Π_2 with probability y and Π_1 with probability $1 - y$. We have $0 < y < x < 1$.

Now consider the following set-up. First, the owners of the firm choose a contract (an incentive scheme) for the manager. At this stage, they do not yet know what the profit of the firm will turn out to be. The contract therefore specifies the wage of the manager for each of the two possible values of Π. The objective of the owners is to maximise expected net profit $E(\Pi - w)$. Note that this objective function implies that the owners are risk neutral.

Given the incentive scheme chosen by the owners, the manager decides whether to accept the job or not and, if he accepts, chooses the level of effort that maximises his expected utility $E(U)$. We assume that he can always obtain a reservation wage w_0, and hence utility $U_0 = u(w_0)$, by working outside the firm, so he will never accept to work for the firm if his expected utility from doing so is less than U_0. After the manager has made his choice, the profit is observed and the manager gets paid. The question is what incentive scheme the owners should choose to maximise $E(\Pi - w)$.

If the owners could observe the manager's effort level, there would be no need for an incentive mechanism, since the owners could then impose an effort level on the manager.[3] All that they would need to do is ensure that the manager accepts the job. This implies ensuring that the manager obtains utility exactly U_0: any payment giving him a higher utility would be unnecessary and would reduce the expected profit of the firm. If the owners wanted no effort, they should pay the manager the reservation wage w_0 whatever the profit turned out to be. Faced with this contract, the manager would accept the job and make zero effort. Net expected profit would be $y\Pi_2 + (1 - y)\Pi_1 - w_0$. If the owners wanted high effort, they should pay the manager $w_0 + \Phi$ whatever the profit. The manager would then accept the job and make high effort. Net expected profit would be $x\Pi_2 + (1 - x)\Pi_1 - (w_0 + \Phi)$. Obviously, the owners would choose to impose high effort if and only if:

$$x\Pi_2 + (1 - x)\Pi_1 - (w_0 + \Phi) > y\Pi_2 + (1 - y)\Pi_1 - w_0 \Leftrightarrow (x - y)(\Pi_2 - \Pi_1) > \Phi.$$

Let us assume that this holds, so the owners would prefer high effort, if effort were observable – otherwise the problem under unobservable effort level would be trivial.

Activity

Prove that, if effort is observable, the owners offer w_0 if they want no effort and $w_0 + \Phi$, if they want high effort. Conclude that, if effort is observable, the risk averse party bears no risk.

Answer

Solve the following trivial maximisation problem for the owners: choose w to maximise $E(\Pi - w)$ subject to the manager getting utility at least equal to $U0$. Since the manager gets the same wage whatever the realisation of profit turns out to be, he bears no risk. All the risk is borne by the owners (the risk neutral party).

The need for an incentive scheme arises when the effort level of the manager cannot be observed by the owners and hence cannot be prescribed in the contract. If they want to induce the manager to exert high effort, the owners must reward the manager with a higher wage in the event that profit turns out to be Π_2 rather than Π_1. More specifically, the owners must design a wage structure $w_i(\Pi_i)$, $i = 1, 2$, that maximises their expected net profit:

$x(\Pi_2 - w_2) + (1 - x)(\Pi_1 - w_1)$

subject to ensuring that the manager accepts the job and chooses to exert high effort, that is subject to a 'participation constraint':

$xu(w_2 - \Phi) + (1 - x)u(w_1 - \Phi) \geq u(w_0)$

and an 'incentive-compatibility constraint':

$xu(w_2 - \Phi) + (1 - x)u(w_1 - \Phi) \geq yu(w_2) + (1 - y)u(w_1).$

The first constraint says that, under the incentive scheme, the expected utility of the manager if he exerts high effort is at least U_0, so the manager will accept the job. The second constraint says that the expected utility of the manager if he makes high effort (the left-hand side of the inequality) is at least as large as his expected utility if he makes zero effort (the right-hand side of the inequality), so the manager chooses to make high effort.

It turns out that in this maximisation problem both constraints are satisfied with equality. The incentive scheme chosen by the owners, if they want to induce high effort, will have the following properties:

- the manager will be rewarded if profit is high: $w_2 > w_1$. This can be derived from the incentive-compatibility constraint

- the expected wage $xw_2 + (1 - x)w_1$ will be higher than $w_0 + \Phi$, the wage under observability of effort. This is a result of the concavity of u. Hence, the owners' net profit will be lower.

Note that if the owners want to induce no effort, all they need to do is offer w_0 whatever the profit; this will ensure that the manager participates and chooses to make no effort. Net profit will be the same as under observability of effort. What will the owners choose to do, offer a contract that induces high effort or one that induces no effort? It depends on whether their maximised net profit is higher under high effort or under zero effort. We have assumed, of course, all along that $(x - y)(\Pi_2 - \Pi_1) > \Phi$, that is to say the owners would prefer high effort to zero effort, if effort were observable. But this does not ensure that the same is true when effort is unobservable, because unobservability reduces the owners' net profit

under high effort but not under no effort. In other words, an additional effect of unobservability is that the owners are more likely to tolerate managerial slack.

Limits to managerial discretion

There are several other mechanisms, apart from direct monetary incentives of the kind examined above, that can limit managerial discretion and reduce slack. Some of the most important are as follows:

- **The threat of takeovers**. The idea is that if managers fail to maximise profits, the stock market value of the firm will be lower, and this will induce outsiders to take over the firm and replace the managers. The effectiveness of this mechanism is reduced by the fact that the collection of information on the firm by outsiders may be costly, takeovers may be subject to free-rider problems, and managers may resist takeovers. Takeovers may also have perverse incentive effects, for instance they may cause managers to put too much emphasis on short-term profits to the detriment of long-term profits.

- **Reputation effects**. Managers care for their careers and are eager to acquire good reputations. This may reduce slack, and may even cause managers to work too hard (i.e. harder than the socially optimal level) early in their career.

- **Supervision**. Monitoring the managers' (more generally, the employees') performance in order to obtain better information on their 'effort' may be costly but feasible. The effectiveness of this mechanism may be reduced by the difficulty of measuring individual effort when team work is important, and also by the possibility of collusion between supervisors and supervisees.

- **Competition in the product market**. The effects of product market competition on managerial incentives can take several forms and may sometimes be ambiguous.[4] One idea is that if a firm does not maximise profits, there is a higher probability that it will not be able to compete with more efficient firms and will therefore go bankrupt. Managers wishing to avoid this will work hard to maximise profits.

- **Organisational form**. The internal organisation of a firm can help mitigate managerial slack, especially by lower managers. The 'unitary form' firm allows greater specialisation of labour, but supervision by the top management becomes more difficult as the firm grows. In the 'multi-divisional form' firm, on the other hand, it is possible for the top management to measure the performance of the different divisions within the firm and compare them with one another. One reason for the gradual decline of the U-form and the emergence of the M-form may have been the need to limit managerial discretion.

[4] A review of the literature is given in Nickell (1996).

Empirical evidence

Empirical work on the performance of firms (for instance, Nickell 1996, Nickell et al. 1996) has looked at a number of factors external to the firm that are associated with improved productivity growth in UK firms. Three such factors have been identified: product market competition, financial market pressure (i.e. a high level of debt) and shareholder control (i.e. the existence of a dominant external shareholder from the financial sector). Using industry-level data, Symeonidis (2008) has found clear evidence of a negative effect of cartels on productivity. Other studies have established a positive effect of trade liberalisation on the productivity of firms in various countries.

The profit-maximisation hypothesis

Although there are many ways in which managerial discretion can be restricted, none of them is perfect, so we should expect deviations from profit maximisation due to the separation of ownership and control. In addition, performing complex calculations is a time-consuming, effort-demanding, and sometimes impossible task, especially under conditions of uncertainty, so members of a firm may often follow simple 'rules of thumb'. Ultimately, the question is how significant any deviations from profit maximisation are likely to be. If we accept, as many economists do, that large deviations from profit maximisation will not allow a firm to survive in the long run, then the hypothesis that firms maximise expected profits seems a reasonable approximation of firm behaviour for most purposes, and in particular for the analysis of the interaction between firms in the market, which is the main focus of 99 Industrial economics.

A reminder of your learning outcomes

By the end of this chapter and having completed the Essential reading and Activities, you should be able to:

- explain the implications of the separation of ownership and control in modern large companies
- analyse optimal incentive mechanisms offered by the owners of a firm to the firm's manager
- describe different mechanisms that may restrict managerial discretion and discuss their limitations
- assess the validity of the profit-maximisation hypothesis.

Sample examination questions

The profit of a firm can take one of two values, Π_1 and Π_2, where $\Pi_2 - \Pi_1 > 10$. The firm is run by a manager who chooses between two levels of effort, $e = 1$ (high) and $e = 0$ (low). The manager's utility function is $U = w^{1/2} - e$, where w is her wage. Whether the firm makes Π_1 or Π_2 depends on the manager's effort and on the firm's environment, which is uncertain. In particular, if the manager makes high effort, the profit is Π_2 with probability 0.8 and Π_1 with probability 0.2. If the manager makes no effort, the profit is Π_2 with probability 0.3 and Π_1 with probability 0.7. Before the realisation of Π is observed, the owners of the firm choose a contract for the manager which specifies the value of w for each of the two possible values of Π. The owners' objective is to maximise expected net profit $E(\Pi - w)$. Given the incentive scheme chosen by the owners, the manager decides whether to take the job and, if she accepts, chooses e to maximise her expected utility $E(U)$. Her reservation wage is $w_0 = 4$. After the manager has made her choice, the profit is observed and the manager gets paid.

a. What is the optimal contract if the owners can observe the manager's effort?

b. What is the optimal contract if the owners cannot observe the manager's effort?

c. Show that the net profit of the owners is lower if the manager's effort is unobservable than if it is observable.

2. 'An optimal incentive scheme offered by the owners of a firm to the firm's manager should reward the manager when profits are high and penalise him when profits are low.' Discuss this statement with reference to an economic analysis of the relationship between the owners and the manager that takes into account the fact that the manager's effort level may not be observable by the owners.

Chapter 3: Short-run price competition

Learning outcomes

By the end of this chapter and having completed the Essential reading and Activities, you should be able to:

- describe and derive the Bertrand paradox

- analyse how the introduction of capacity constraints in the Bertrand model leads to equilibrium outcomes with price greater than marginal cost and positive profits

- explain the theoretical foundations of the Cournot model

- analyse the Cournot model for various assumptions regarding the demand, the number of firms, and the cost structures.

Essential reading

Church, J.R. and R. Ware *Industrial Organization: A Strategic Approach*. Chapter 8.

Tirole, J. *The Theory of Industrial Organization.* Chapter 5.

Further reading

Books

Cabral, L. *Introduction to Industrial Organization*. (Cambridge, MA: MIT Press, 2000). Chapter 7.

Carlton, D.W. and J.M. Perloff *Modern Industrial Organization*. (Pearson Addison Wesley, 2005) fourth edition. Chapter 6.

Martin, S. *Industrial Economics*. (Englewood Cliffs, NJ: Prentice Hall, 1994) second edition. Chapter 5.

Shy, O. *Industrial Organization*. (Cambridge, MA: MIT Press, 1995). Chapter 6.

Introduction

This chapter begins our analysis of firms' conduct in oligopolistic markets. We will start with the simplest strategic situations. In particular, we will assume that the only decision firms have to make is to set a price for their product, or a level of output. In fact, of course, firms can use many instruments to compete in a market, and subsequent chapters will analyse several examples of more complex strategic situations. The main reason why it makes sense to abstract from these additional considerations in this chapter is that price is an instrument that firms can change relatively easily in the short run. On the other hand, other instruments are more difficult to change. These include product design, the level of capacity, an advertising-based brand image, product quality or cost determined by research and development (R&D), and so on; ultimately there is also the decision of whether or not to enter or stay in a market. Since these long-run decisions are relatively difficult to change, they are taken as given when making shorter-term decisions. Thus we can proceed to analyse short-run competition between firms in the context of fixed cost structures and product characteristics, and with a fixed number of firms in the market.

In this chapter we focus on **static** models of oligopoly: the firms interact only once in the market and their actions are simultaneous. Repeated interaction is the subject of the next chapter.

Throughout Part 2 of the guide, we will be using the theory of non-co-operative games to model firms' interaction. The equilibrium market outcomes are therefore the equilibria of these games. To identify them, we will be making use of some elementary solution concepts from non-co-operative game theory: the Nash equilibrium and the subgame-perfect equilibrium.[1]

[1] *See Tirole (1988), pp.423–432, or Shy (1995), Chapter 2, for an introduction to these game-theoretic concepts and discussion of applications.*

The basic Bertrand model

Consider a very simple set up as follows. There are two firms, 1 and 2, producing a homogeneous product (the result easily generalises to N firms). The two firms interact only once and they simultaneously and independently set prices p_1 and p_2 respectively. The market demand for the product is given by $q = D(p)$, and both firms have the same constant marginal cost c. The firm with the lowest price gets all the market demand at that price; if the two prices are the same each firm gets half the market demand at that price.

The Nash equilibrium outcome of this game is $p_1{}^* = p_2{}^* = c$. In other words, firms price at marginal cost and make zero profit.

Activity

Prove this result.

Answer

The proof consists in distinguishing cases and showing that in all of them except the case $p_1 = p_2 = c$ there exists a profitable deviation by at least one firm. Read, for example, Tirole (1988), p. 210.

The intuition is that unless prices are the same and equal to c, each firm has an incentive to undercut the other. Note that there is a qualification to this result for the case of asymmetric marginal costs: in that case, at equilibrium the low-cost firm sets a price marginally lower than the cost of the high-cost firm[2] and makes positive profit. Still, this profit is small if the cost difference is small, and the high-cost firm makes no sales and no profit. The outcome of the simple Bertrand game has therefore justifiably been called the 'Bertrand paradox'.

[2] *Provided this is not higher than the monopoly price corresponding to its own cost; otherwise it sets the monopoly price.*

There are three resolutions to this paradox: repeated interaction, product differentiation and capacity constraints.

Bertrand competition with capacity constraints

To understand why capacity constraints matter, take the simple model of the previous section and assume that both firms have production capacity smaller than $D(c)$, that is to say no firm can cover the entire demand at a price equal to the common marginal cost. Then $p_1 = p_2 = c$ is no longer a Nash equilibrium. Why? If firm i raises its price slightly above c, given that $p_j = c$, all consumers will want to buy from firm j; however, firm j will not be able to satisfy the whole demand, so some consumers will end up buying from firm i. Hence firm i will make positive profit instead of zero. Since $p_i = c$ is not a profit-maximising response to $p_j = c$, $p_1 = p_2 = c$ is not a Nash equilibrium.

The exact equilibrium outcome in the above model depends on what specific assumption we make about the way consumers are rationed. In general, however, models with capacity constraints have Nash equilibria with price greater than marginal cost and positive profits. Note that rigid

capacity constraints are a special case of decreasing returns to scale (i.e. a technology such that the marginal cost increases with output). Such models also have equilibria with price greater than marginal cost.

We now look at a simple model to fix these ideas.[3] An additional important implication of this model is that, in certain cases, a game where capacity-constrained firms compete in prices is formally equivalent to a game where firms set quantities and an auctioneer determines the market-clearing price.

Consider a market where two firms, 1 and 2, produce a homogeneous product. Demand is $q = D(p) = 1 - p$, or equivalently $p = 1 - q_1 - q_2$. The firms have capacity constraints $\bar{q}_1$ and $\bar{q}_2$, and we assume that $\bar{q}_i < 1/3$, $i = 1, 2$. The marginal cost of production is zero for $q_i < \bar{q}_i$ and infinite for $q_i > \bar{q}_i$. Finally, we assume that consumers are rationed according to the 'efficient' rationing rule. The rationing of consumers results from the fact that the low-price firm cannot serve the entire market. The question then arises as to which consumers end up buying from the high-price firm; this is important because it determines the shape of the residual demand of the high-price firm. For instance, suppose that firm 1 is the low-price firm. Then under the efficient rationing rule, the residual demand of firm 2 is given by $D(p_2) - \bar{q}_1$ if $D(p_2) > \bar{q}_1$, and zero otherwise. This is illustrated in Figure 3.1, which depicts both the market demand $D(p)$ and the residual demand of firm 2 (note that with price p_1 the low-price firm 1 sells up to capacity).
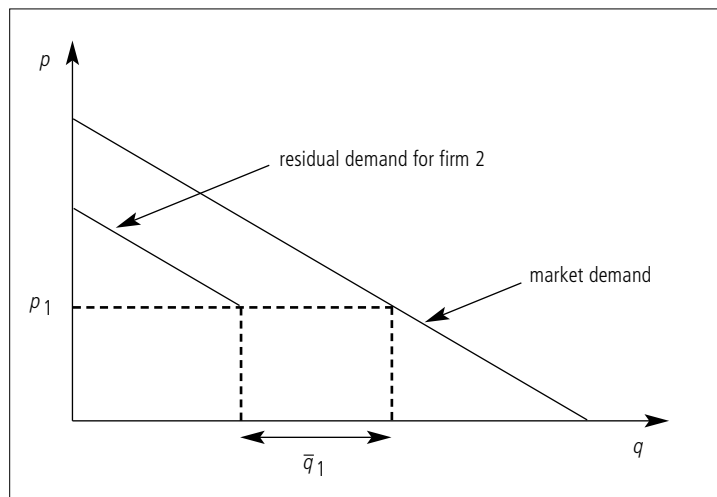
Figure 3.1

We will now show that the unique Nash equilibrium outcome of this game is for both firms to set the price $p^* = 1 - \bar{q}_1 - \bar{q}_2$. At this price both firms sell up to their respective capacities and the market clears. Note that this price is higher than marginal cost (which is zero), and therefore implies positive profits for both firms.

To show that this is a Nash equilibrium, we need to show that none of the firms has an incentive to unilaterally deviate from this equilibrium. Is it profitable for firm i to set a price lower than p^*, given that firm j sets price p^*? The answer is no. By charging p^* firm i sells exactly $\bar{q}_i$. Now firm i cannot produce more than $\bar{q}_i$ anyway, so by reducing its price below p^* it would simply sell the same quantity at a lower price and would therefore make less profit.

Is it profitable for firm i to set a price higher than p^*, given that firm j sets price p^*? The answer is again no, but the argument now is slightly more subtle. Suppose that firm i sets a price $p \geq p^*$. Then it has residual demand

$1 - p - \overline{q}_j$, because at price p total market demand is given by $1 - p$ and firm j sells $\overline{q}_j$. Firm i makes profit $\Pi = p(1 - p - \overline{q}_j)$. Using the inverse demand function, the expression for profit can be written as $(1 - q - \overline{q}_j)q$, where q is the quantity sold by firm i at price p. Note that the profit function $\Pi = (1 - q - \overline{q}_j)q$ is exactly the same as the profit function of a firm that chooses output q given that the rival firm chooses output $\overline{q}_j$. This profit function is concave in q, that is $\Pi''(q) < 0$. Also, $\partial\Pi/\partial q = 1 - 2q - \overline{q}_j$. Evaluated at $q = \overline{q}_i$ this derivative is equal to $1 - 2\overline{q}_i - \overline{q}_j$, which is positive because both $\overline{q}_i$ and $\overline{q}_j$ are less than 1/3. In other words, if firm i starts from $\overline{q}_i$ and reduces its quantity, its profit will fall.[4] This is another way of saying that if firm i starts from p^* and increases its price, its profit will fall.

We have therefore shown that it is not profitable for firm i either to set a price lower than p^* or to set a price higher than p^*, given that firm j sets price p^*. Hence $p_1 = p_2 = p^*$ is the unique Nash equilibrium of the game.

There are two main conclusions of the above analysis. First, we have seen how a static pricing game between symmetric firms can lead to $\Pi > 0$ if there are capacity constraints. Second, everything is as if firms put quantities equal to their capacities on the market and an auctioneer determines the price that clears the market. The equilibrium of this game is the Cournot equilibrium (see the analysis of the Cournot model below).

Choice of capacities

One issue which was swept under the carpet in the above discussion is the choice of capacities. We assumed that firms were capacity constrained, and significantly so ($\overline{q}_i < 1/3$, i = 1, 2). But can't firms build capacities that would allow them to cut price down to marginal cost and supply the whole market if they so choose? To examine this question, we would need to construct a more complex game than the one we have analysed, namely a two-stage game with choice of capacities in the first stage and price competition in the second stage, when capacities are taken as fixed. Now intuitively one would expect firms to strategically refrain from building too much capacity because this would destroy their profits in the price competition stage. This is exactly what the formal analysis of such games predicts. In fact there is a much stronger result, due to Kreps and Scheinkman: if demand is concave and the rationing rule is the efficient one, then the outcome of this two-stage game is the same as the outcome of the one-stage Cournot game (which involves $\Pi > 0$, as we will see below).

One difficulty in oligopoly theory has been that the widely used Cournot model, which assumes that firms compete by setting quantities, may lack strong foundations, since firms typically compete by setting prices, not quantities. We have seen, however, that the Cournot model can be interpreted in either of the following ways:

- as a one-stage pricing game between capacity-constrained firms

- as a reduced-form game for the two-stage game with choice of capacities in the first stage and price setting in the second stage.

Of course, both these results rest on particular assumptions concerning the rationing rule, so in more general settings we would not get equilibrium outcomes that look exactly like the Cournot outcome. However, it is generally valid to think of quantity competition as a choice of capacity or scale that determines the firms' cost function and hence the conditions of price competition. It is therefore valid to interpret the distinction between price competition (Bertrand) and quantity competition (Cournot) as a

[4] *The concavity of the profit function ensures that any reduction of q below $\overline{q}_i$ will reduce profit.*

difference in the flexibility of production: if costs rise steeply with output in the short run in a particular industry, then the Cournot model is more appropriate for this industry; if not, then the Bertrand model is more appropriate.

The Cournot model

Consider a model of competition between two firms, 1 and 2, producing a homogeneous product. The inverse demand function has the general form $p = P(q_1 + q_2)$, where q_1 and q_2 are quantities produced by firm 1 and firm 2 respectively and p is the market price. The demand curve is downward sloping, so $P'(q_1 + q_2) < 0$. The total cost of firm i is given by $C_i(q_i)$, $i = 1$, 2. The two firms meet only once and they simultaneously set quantities.

The Nash equilibrium of this game is computed as follows. Firm 1 chooses q_1 to maximise its profit $\Pi 1 = q_1 P(q_1 + q_2) - C_1(q_1)$, taking q_2 as given. The first-order condition for this maximisation problem is:

$$\frac{\partial \Pi^1}{\partial q_1} = 0 \Leftrightarrow P(q_1 + q_2) - C_1'(q_1) + q_1 P'(q_1 + q_2) = 0$$

This equation implicitly defines the optimal choice of q_1 for any given level of q_2. It is called the 'reaction function' of firm 1.

Similarly for firm 2. That is, it chooses q_2 to maximise $\Pi 2 = q_2 P(q_1 + q_2) - C_2(q_2)$, taking q_1 as given. The first-order condition, or the reaction function of firm 2, is:

$$\frac{\partial \Pi^2}{\partial q_2} = 0 \Leftrightarrow P(q_1 + q_2) - C_2'(q_2) + q_2 P'(q_1 + q_2) = 0$$

The Nash equilibrium is the solution of the system of the two first-order conditions.

We can identify some interesting properties of this equilibrium just by looking at the first order conditions:

- The price is greater than marginal cost, namely:
  $p - C_i' = -q_i P'(q_1 + q_2) > 0$, $i = 1, 2$.

- The price is lower than the monopoly price. To see this note that the first-order condition for a monopolist would be similar to the above first-order conditions except that the third term would be $(q_1 + q_2)P'(q_1 + q_2)$. Hence the difference between price and marginal cost would be greater under a monopolist.

- The first-order condition for firm i can be written as $(p - C_i')/p = s_i/\varepsilon$, where $s_i = q_i/(q_i + q_j)$ is the market share of firm i and ε is the absolute value of the elasticity of demand: $1/\varepsilon = -(1/p)(q_1 + q_2)[P'(q_1 + q_2)]$. Thus the price-cost margin $(p - C_1')/p$ (also called the 'Lerner index') increases with the market share of firm i and decreases with the elasticity of demand.

Activity

Derive the result $(p - C_1')/p = s_1/\varepsilon$ from the first-order condition for firm 1.

Answer

Straightforward algebraic manipulations yield the result.

To derive the equilibrium explicitly, let us further assume a linear demand function $q = a - p$ and constant marginal costs c_1, c_2. Then:

$$\frac{\partial \Pi^1}{\partial q_1} = a - 2q_1 - q_2 - c_1 = 0 \Leftrightarrow q_1 = R_1(q_2) = \frac{a - q_2 - c_1}{2}$$

and:

$$\frac{\partial \Pi^2}{\partial q_2} = a - 2q_2 - q_1 - c_2 = 0 \Leftrightarrow q_2 = R_2(q_1) = \frac{a - q_1 - c_2}{2}$$

The functions R_1 and R_2 are the reaction functions of firm 1 and firm 2 respectively. Solving the system of the two equations we obtain the equilibrium quantities:

$$q_1^* = \frac{a - 2c_1 + c_2}{3} \qquad q_2^* = \frac{a - 2c_2 + c_1}{3}$$

Thus a firm's output decreases with own marginal cost and increases with rival marginal cost: the more efficient a firm the higher its market share. A firm's output also increases in the exogenous demand shift parameter a. It can be checked that the same is true for the equilibrium profits. These are positive even when the firms are symmetric ($c_1 = c_2$), a result which is in sharp contrast with the Bertrand model.

Activity

Derive the equilibrium profits.

Answer

Substitute q_1^* and q_2^* into the profit functions $\Pi^1 = q_1(p - c_1)$ and $\Pi^2 = q_2(p - c_2)$ and use the inverse demand function $p = a - q_1 - q_2$. You should obtain $\Pi^{1*} = (a - 2c_1 + c_2)^2/9$, $\Pi^{2*} = (a - 2c_2 + c_1)^2/9$.

The analysis proceeds along similar lines when the number of firms is greater than two. An interesting property of the Cournot model with N firms is that, under certain conditions regarding demand and costs, total industry profits increase as concentration in an industry rises.[5] This may be taken as one justification for the view that higher concentration increases prices and profits because firms have more market power. However, bear in mind that concentration is itself endogenous, and that both concentration and profitability are ultimately determined by basic industry characteristics, such as technology and demand. As discussed in Chapter 9 of this guide, a positive association between the two need not imply a causal link or may not exist at all once the endogeneity of concentration is taken into account.

[5] See Tirole (1988), pp.218–223.

Because profits are higher under quantity-setting (Cournot) than under price-setting (Bertrand), the two models are often interpreted as representing different degrees of competition. This interpretation is fine for games where firms do not make any long-run decisions, except perhaps the decision to enter or not the market. It is not appropriate when firms make long-run choices such as investment, advertising, R&D, etc. before setting prices or quantities.[6]

[6] The reason will become clear when you have read Chapter 5 of this guide.

Activities

1. Consider the model of price competition with capacity-constrained firms analysed above. Assume, however, that one of the firms, say firm 1, has capacity higher than 1/3. In particular assume that $\overline{q}_1 = 2/5$. Under what condition is $p_1 = p_2 = p^* = 1 - \overline{q}_1 - \overline{q}_2$ still the unique Nash equilibrium of the game?

2. Consider a market where N symmetric firms produce a homogeneous product and compete by simultaneously setting quantities. The inverse demand function is given by $p = a - Q$, where Q is total quantity produced, that is $Q = q_1 + q_2 + \ldots + q_N$. The marginal cost is constant and equal to c for all firms.

   a. Derive the Cournot-Nash equilibrium. (Hint: After deriving the first order condition for firm i, use the symmetry of the model to significantly simplify the computations.)

   b. Derive the equilibrium price, profit for firm i and industry profit, and show that all three are decreasing in the number of firms N.

3. One way to test the predictions of oligopoly models is to conduct laboratory experiments. Of course, the conditions in the laboratory are very different from the conditions that firms face in the real world. For one thing, players in laboratory experiments are not very experienced and the stakes are much lower. Nevertheless, there is a large literature on laboratory experiments in industrial economics. What have we learned from it? Do agents behave as economic theory predicts? A useful reference is the survey by C.A. Holt, 'Industrial Organization: A Survey of Laboratory Research', published in Kagel, J. and A. Roth (eds), *Handbook of Experimental Economics* (Princeton University Press, 1995). It is also available online at: http://people.virginia.edu/~cah2k/iosurvtr.pdf

   Another useful reference is the January 2000 special issue of the *International Journal of Industrial Organization* on 'Experimental Economics and Industrial Organization'. Both these references cover a range of topics, from simple static games to more complicated dynamic games with commitment, product differentiation, asymmetric information, and so on.

A reminder of your learning outcomes

By the end of this chapter, and having completed the Essential reading and Activities, you should be able to:

- describe and derive the Bertrand paradox

- analyse how the introduction of capacity constraints in the Bertrand model leads to equilibrium outcomes with price greater than marginal cost and positive profits

- explain the theoretical foundations of the Cournot model

- analyse the Cournot model for various assumptions regarding the demand, the number of firms, and the cost structures.

Sample examination questions

1. a. Describe and prove the Bertrand paradox.

   b. Consider a market with two price-setting firms producing a homogeneous product. The demand function is $q = D(p) = 1 - p$, which implies the inverse demand $p = 1 - q_1 - q_2$. The two firms have capacity constraints $\overline{q}_1$ and $\overline{q}_2$, where $\overline{q}_1 + \overline{q}_2 = 3/5$. The marginal cost of production is zero for $q_i \le \overline{q}_i$ and infinite for $q_i > \overline{q}_i$. Finally, assume that consumers are rationed according to the efficient rationing rule.

      i. Show that if $\overline{q}_1 = \overline{q}_2$, there is a unique Bertrand-Nash equilibrium where $p_1 = p_2 = p^* = 1 - \overline{q}_1 - \overline{q}_2$

      ii. Show that when $\overline{q}_1 \ne \overline{q}_2$, the equilibrium under part (i) breaks down when the firms' capacities are too dissimilar.

2. a. Consider a market with two quantity-setting firms producing a homogeneous product. The inverse demand function is given by $p = 1 - q_1 - q_2$ and the two firms have constant marginal costs c_1 and c_2 such that $c_1 + c_2 = 2c$, where c is a constant.

   i. Compute the Cournot-Nash equilibrium.

   ii. Show that as the two firms become more asymmetric (i.e. c_i moves away from c), total industry profit increases.

   iii. Compute an index of concentration in this market and show that it increases as the two firms become more asymmetric.

   b. A researcher has estimated a model of industry profitability using cross-industry data and has found a positive coefficient on the concentration variable. He claims that the results show that higher concentration leads to higher industry profit. What is the theoretical basis for this claim? Do you agree with this conclusion? What would your advice be to a policy-maker worried about the high level of concentration in many industries?