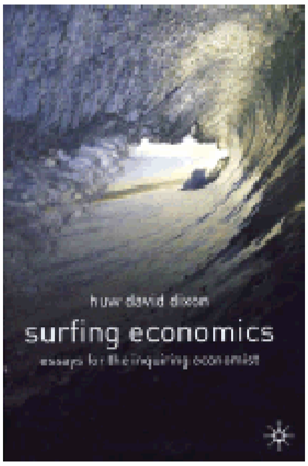
Oligopoly Theory Made Simple



Chapter 6, *Surfing Economics*, pp 125-160.

Chapter 6. Oligopoly Theory Made Simple

6.1 Introduction.

Oligopoly theory lies at the heart of industrial organisation (IO) since its object of study is the interdependence of firms. Much of traditional micro-economics presumes that firms act as passive price-takers, and thus avoids the complex issues involved in understanding firms' behaviour in an interdependent environment. As such, recent developments in oligopoly theory cover most or all areas of theoretical IO, and particularly the "new" IO. This survey is therefore very selective in the material it surveys: the goal is to present some of the basic or "core" results of oligopoly theory that thave a general relevance to IO.

The recent development of oligopoly theory is inextricably bound up with developments in abstract game theory. New results in game theory have often been applied first in the area of oligopoly (for example, the application of mixed strategies in the 1950s – see Shubik 1959, and more recently the use of subgame perfection to model credibility). The flow is often in the opposite direction: most recently, the development of sequential equilibria by Kreps, Milgrom, Roberts, and Wilson arose out of modelling reputational effects in oligopoly markets. Over recent years, with the new IO, the relationship with game theory has become closer. This chapter therefore opens with a review of the basic equilibrium concepts employed in the IO – Nash equilibrium, perfect equilibrium, and sequential equilibrium.

The basic methodology of the new IO is neo-classical: oligopolistic rivalry is studied from an equilibrium perspective, with maximising firms, and uncertainty is dealt with by expected profit of payoff maximization. However, the subject matter of the new IO differs significantly from the neo-classical micro-economics of the standard textbook. Most importantly, much of the new IO focuses on the process of competition over time, and on the effects of imperfect information and uncertainty. As such, it has expanded its vision from static models to consider aspects of phenomena which Austrian economists have long been emphasising, albeit with a rather different methodology.

The outline of the chapter is as follows. After describing the basic equilibrium concepts in an abstract manner in the first section, the subsequent two sections explore and contrast the two basic static equilibria employed by oligopoly theory to model product market competition – Bertrand (price) competition, and Cournot (quantity) competition. These two approaches yield very different results in terms of the degree of competition, the nature of the first-mover advantage, and the relationship between market structure (concentration) and the price-cost margin.

The fourth section moves on to consider the incentive of firms to precommit themselves in sequential models; how firms can use irreversible decisions such as investment or choice of managers to influence the market outcome in their favour. This approach employs the notion of subgame perfect equilibria, and can shed light on such issues as whether or not oligopolists will overinvest, and why non-profit maximizing managers might be more profitable for their firm than profit maximizers. The fifth section explores competition over time, and focuses on the results that have been obtained in game-theoretic literature on repeated games with perfect and imperfect information. This analysis centres on the extent to which collusive outcomes can be supported over time by credible threats, and the influence of imperfect information on a firm's behaviour in such a situation. Alas, many areas of equal interest have had to be omitted – notably the literature on product differentiation, advertising, information transmission, and price wars. References are given for these in the final section.

Lastly, a word on style. I have made the exposition of this chapter as simple as possible. Throughout the chapter I employ a simple linearized model as an example to illustrate the mechanics of the ideas introduced. I hope that readers will find this useful, and I believe that it is a vital complement to general conceptual understanding. For those readers who appreciate a more rigorous and general mathematical exposition, I apologise in advance for what may seem sloppy in places. I believe, however, that many of the basic concepts of oligopoly theory are sufficiently clear to be understood without a general analysis, and that they deserve a wider audience than a more formal exposition would receive.

6.2 Non-cooperative equilibrium

The basic equilibrium concept employed most commonly in oligopoly theory is that of the *Nash equilibrium*, which originated in Cournot's analysis of duopoly (1838). The Nash equilibrium applies best in situations of a one-off game with perfect information. However, if firms compete repeatedly over time, or have imperfect information, then the basic equilibrium concept needs to be refined. Two commonly used equilibrium concepts in repeated games are those of *subgame perfection* (Selten 1965), and if information is imperfect, *sequential equilibria* (Kreps *et al.* 1982).

We shall first introduce the idea of a Nash equilibrium formally, using some of the terminology of game theory. There are n firms, i = 1, who each choose some strategy a_i from a set of feasible actions A_i. The firm's strategy might be one variable (price/quantity/R&D) or a *vector* of variables. For simplicity, we will take the case where each firm chooses one variable only. We can summarize what each and every firm does by the n-vector $(a_1, a_2 \ldots, a_n)$. The "payoff" function shows the firm's profits p_i as a function of the strategies of each firm:

$$p_i = p_i(a_1, a_2, \ldots, a_n) \tag{1}$$

The payoff function essentially describes the market environment in which the firms operate, and will embody all the relevant information about demand, costs and so on. What will happen in this market? A Nash equilibrium is one possibility, and is based on the idea that firms choose their strategies non-cooperatively. A Nash equilibrium occurs when each firm is choosing its strategy optimally, given the strategies of the other firms. Formally, the Nash equilibrium is an n-vector of strategies $(a_1^*, a_2^* \ldots, a_n^*)$ such that for each firm i, a_i^* yields maximum profits given the strategies of the n – 1 other firms a_{-i}^*.[1] That is, for each firm:

$$p_i(a_i^*, a_{-i}^*) \geq p_i(a_i, a_{-i}^*) \tag{2}$$

for all feasible strategies $a_i \in A_i$. The Nash equilibrium is often defined using the concept of a *reaction function*. A reaction function for firm i gives its best response given what the other forms are doing. In a Nash equilibrium, each firm will be on its reaction function.

Why is the Nash equilibrium so commonly employed in oligopoly theory? Firstly, because no firm acting on its own has any incentive to deviate from the equilibrium. Secondly, if all firms expect a Nash equilibrium to occur, they will choose their Nash equilibrium strategy, since this is their best response to what they expect the other firms to do. Only a Nash equilibrium can be consistent with this rational anticipation on the part of firms. Of course, a Nash equilibrium may not exist, and there may be multiple equilibria. There are many results in game theory relating to the existence of Nash equilibrium. For the purpose of inductrial economics, however, perhaps the most relevant is that if the payoff functions are continuous and strictly concave in each firm's own strategy then at least one equilibrium exists.[2] Uniqueness is rather harder to ensure, although industrial economists usually make strong enough assumptions to ensure uniqueness.[3]

If market competition is seen as occurring over time, it may be inappropriate to employ a one-shot model as above. In a *repeated game* the one-shot *constituent* game is repeated over T periods (where T may be finite or infinite). Rather than simply choosing a one-off action, firms will choose an action a_{it} in each period $t = 1\ldots, T$. For repeated games, the most commonly used equilibrium concept in recent oligopoly theory literature is that of subgame perfection which was first formalised by Selten (1965), although the idea had been used informally (e.g. Cyert and De Groot 1970). At each time t, the firm will decide on its action a_{it} given the past history of the market h_t, which will include the previous moves by all firms in the market.

A firm's "strategy" in the repeated game[4] is simply a rule s_i which the firm adopts to choose its action a_{it} at each period given the history of the market up to then, h_t:

$$a_{it} = s_i(h_t)$$

If we employ the standard Nash equilibrium approach, an equilibrium in the repeated game is simply n strategies$(s_1^*, s_2^* ..., s_n^*)$ such that each firm's strategy s_i^* is optimal given the other firms' strategies s_{-i}^*. Thus no firm can improve its payoff by choosing a different strategy, given the strategies of the other firms.

However, a major criticism of using the standard Nash equilibrium in repeated games is that it allows firms to make "threats" which are not *credible*, in the sense that it would not be in their interest to carry out the threat. For example, consider the example of entry deterrence, with two periods. In the first period, the entrant decides whether or not to enter. In the second period, the entrant and incumbent choose outputs. The incumbent could adopt the following strategy: if entry does not occur, produce the monopoly output. If entry does occur, produce a very large output which drives down the price below costs at whatever output is chosen by the entrant. In effect, the entrant is posed with a powerful threat by the incumbent: "if you enter, I'll flood the market and we'll both lose money". Clearly, with this powerful threat, the incumbent will be able to deter any entry. However, it is not a *credible* threat: if entry *were* to occur, then the incumbent would not wish to carry out this potent threat. Thus the incumbent's strategy is not credible, since he would be making unnecessary losses.

Subgame perfection was formulated to restrict firms to *credible* strategies. The basic idea of subgame perfection is quite simple. In a Nash equilibrium the firm chooses its strategy s_i "once and for all" at the beginning of the game, and is committed to it throughout the play (as in the above example). To rule out non-credible threats, however, in a subgame-perfect equilibrium, at each point in time firms choose their strategy for the rest of the game. The "subgame" at any time t is simply the remainder of the game from t through to the last period T. Subgame perfection requires that the strategies chosen are Nash equilibria in each subgame. This rules out non-credible threats, since in effect it requires a firm to choose its strategy optimally at each stage in the game. In our example, the incumbent's threat to expand output is not "credible": in the subgame consisting of the second period, it is not a Nash equilibrium. Indeed, if the market is Cournot and there is a unique Cournot equilibrium, then the unique subgame-perfect strategy for the incumbent involves producing the Cournot output if entry has occurred. One of the major attractions of subgame perfection is that it narrows down the number of equilibria: there are often multiple Nash equilibria in repeated games and imposing "credibility" on strategies can reduce the number considerably, at least in finitely repeated games.

With imperfect information, a commonly used equilibrium concept is that of a "sequential" equilibrium (Kreps *et al*. 1982). This is formally a rather complex concept, but we shall provide a simple example in the section "competition over time". The basic idea of subgame perfection is employed, with the added ingredient of Bayesian updating of information.[5] Firms may be uncertain about each other's payoff functions (e.g. they do not know each other's costs or each other's objectives). At the start of the game, firms have certain prior beliefs, which they then update through the game. Firms may be able to learn something about each other from each other's actions. In such a situation, firms of a certain type may be able to build a "reputation" by taking actions which distinguish themselves from firms of another type. For example, in Milgrom and Robets' (1982b) paper on entry deterrence, low-cost incumbents are able to distinguish themselves from high-cost incumbents by following a "limit pricing" strategy which it is unprofitable for high-cost firms to pursue. These reputational equilibria are very important since they can explain how firms might behave against their short-run interest in order to preserve their reputation intact, for example as a low-cost firm or as an aggressive competitor.

6.3 Cournot and Bertrand equilibria with homogeneous products

The previous section considered the concept of a Nash equilibrium in purely abstract terms. To make the concept concrete we need to specifiy the exact nature of the strategies chosen and define the payoff function. Corporate strategy is, of course, very broad embracing all the activities of the firm – price, output, investment, advertising, R & D and so on. In practice, oligopoly theory abstracts from the complexity of real-life corporate strategy and concentrates on just one or two strategic variables. There are two basic ways of modelling how firms compete in the market. The first takes the view that the firm's strategic variable is its output and originates in Cournot (1838). The second takes the view that the firm's basic strategic variable is price and, originates in the work of Bertrand (1883), Edgeworth (1925) and more recently in models of imperfect competition with product differentiation (e.g. Chamberlin 1933; Dixit and Stiglitz 1977).

As we shall see, whether price or output is the strategy makes a difference to the equilibrium outcome. For example, one of the basic issues of interest to industrial economists is the relationship between concentration and the price-cost margin. The standard notion that higher concentration leads to a higher price-cost margin is based on the Cournot view, and does not hold in the Bertrand framework where there can be a perfectly competitive outcome with two or more firms. The distinction between price and quantity setting in the context of oligopoly is not present in monopoly, where it makes no difference whether the monopolist chooses a price or quantity (the monopolist simply chooses a point on its demand curve). In order to capture the distinction between the Cournot and Bertrand framework in its starkest form, we will first consider the simplest case of homogeneous goods. We will then discuss what arguments there are for choosing between the two competing approaches to modelling product market competition. In the next section we will pursue this fundamental dichotomy further in the context of the more realistic case of differentiated commodities.

6.3.1 Cournot-Nash equilibrium with homogeneous goods

The basic view of the market taken by Cournot was that firms choose their outputs and that the market then “clears” given the total output of firms. There are n firms $i = 1 \ldots, n$, which produce outputs x_i, industry output being $x = \sum_{i=1}^{n} x_i$. We will make the simplest possible assumption about demand and costs:

*A*1 *Industry demand* $P = 1 - \sum_{i=1}^{n} x_i$ (3)

*A*2 *Firm’s costs* $c(x_i) = cx_i$

Equation (3) is called the *inverse industry demand function*. Normally the industry demand curve is seen as arising from the utility-maximizing behaviour of consumers – the market demand curve tells us how much households wish to buy at a given price. The mathematical operation of taking the inverse, as in (3), has important economic implications: it assumes that there can only be one “market” price. Thus firms have no direct control over the price of their output, only an indirect control via the effect that changes in their own output have on the total industry output.

Given A1 – A2, we can define the firm’s payoff function which gives firm i’s profits as a function of the outputs chosen:

$$
\begin{aligned}
p_i(x_1, x_2, \ldots x_n) &= x_i(1 - \sum_{j=1}^{n} x_j) - cx_i \\
&= x_i - x_i^2 - x_i \sum_{j \neq i} x_j - cx_i
\end{aligned}
\tag{4}
$$

Each firm has a reaction function, which gives its profit-maximizing output as a function of the outputs chosen by the other firms. Since firm i treats the output of the other firm j ? i as fixed, the first-order condition for maximizing (4) with respect to x_i is[6]:

$$\frac{\partial p_i}{\partial x_i} = 1 - \sum_{j \neq i} x_j - 2x_i - c = 0$$

Solving this defines the reaction function for firm i:

$$x_i = \frac{1 - \sum_{j \neq i}^{n} x_j - c}{2} \qquad i = 1, \ldots, n \tag{5}$$

Each firm has a similar reaction function, and the Nash equilibrium occurs when each firm is on its reaction function (i.e. choosing its optimal output given the output of other firms). There will be a symmetric and unique Cournot-Nash equilibrium which is obtained by solving the n equations (5) for outputs (which are all equal by symmetry),

$$x_i^* = x^c = \frac{1-c}{n+1} \qquad \textit{Cournot-Nash equilibrium output} \tag{6}$$

which results in equilibrium price:

$$p^c = \frac{1}{n+1} + \frac{n}{n+1}c \tag{7}$$

For example, if $n = 1$ (monopoly) we get the standard monopoly solution. For $n = 2$, $x_c = (1 - c) / 3$, $p^c = 1/3 + (2/3)c$. The price cost margin for each firm is:

$$m^c = \frac{p^c - c}{p^c} = \frac{1-c}{1+nc} \tag{8}$$

There is a clear inverse relationship between the equilibrium price-cost margin and the number of firms. As the number of firms become infinite $(n \rightarrow \infty)$, the price-cost margin tends to its competitive level of *0*: as the number falls to one, it tends to its monopoly level $(1 - c)/(1 + c)$, as predicted in Figure 1.

Figure 1 The price-cost margin and the number of firms in the Cournot-Nash equilibrium

What is the intuition behind this relationship between number of firms and the price-cost margin? Very simply, with more firms each firm's own demand becomes more elastic. With an infinite number of firms the firm's elasticity becomes infinite and hence the firms behave as competitive price-takers. The representative firm's elasticity $?_i$ can be related to the industry elasticity $?$:

$$h = \frac{p}{x}\frac{dx}{dp} \tag{9a}$$

$$h_i = \frac{p}{x_i}\frac{dx_i}{dp} = \frac{x}{x_i}\left(\frac{p}{x}\frac{dx_i}{dp}\right) \qquad (9b)$$

Under the Nash assumption firms treat the other firms' outputs as given and the change in industry output x equals the change in firm i's output. Hence $dx_i/dp = dx/dp$. Of course x_i/x is the i^{th} firm's market share which, for our example, is in equilibrium $1/n$. Hence (9b) becomes:

$$?_i = n.\ ? \qquad (10)$$

In equilibrium each firm's elasticity is equal to n times the industry elasticity of demand. As n gets large, so does $?_i$ leading to approximately "price-taking" behaviour.

6.3.2 Bertrand competition with homogeneous products

In his famous review, Bertrand criticised Cournot's model on several counts. One of these was the reasonable one that firms set prices not quantities: the output sold by the firm is determined by the demand it faces at the price it sets. What is the equilibrium in the market when firms set prices, the Bertrand-Nash equilibrium?

If firms set prices the model is rather more complicated than in the Cournot framework since there can be as many prices in the market as there are firms. In the Cournot framework the inverse industry demand curve implies a single "market" price. In the Bertrand framework each firm directly controls the price at which it sells its output and, in general, the demand for its output will depend on the price set by each firm and the amount that they wish to sell at that price (see Dixon 1987b). However, in the case of a homogeneous product where firms have constant returns to scale, the demand facing each firm is very simple to calculate. Taking the case of duopoly, if both firms set *different* prices then all households will wish to buy from the lower-priced firm which will want to meet all demand (so long as its price covers cost), and the higher-priced firm will sell nothing. If the two firms set the *same* price, then the households are indifferent between buying from either seller and demand will be divided between them (equally, for example). If firms have constant marginal cost there exists a unique Bertrand-Nash equilibrium with tow or more firms where each firm sets its price p_i equal to marginal cost – the competitive equilibrium. This can be shown in three steps:

Step 1. If both firms set different prices then that cannot be an equilibrium. The higher-priced firm will face no demand and hence can increase profits by undercutting the lower-priced firm so long as the lower-priced firm charges in excess of c. If the *lower*-priced firm charged c, it could increase profits from 0 by raising its price slightly while undercutting the higher-priced firm. Hence any Bertrand-Nash equilibrium must be a single-price equilibrium (SPE).

Step 2. The only SPE is where all firms set the competitive price. If both firms set a price *above* c, then either firm can gain by undercutting the other by a small amount. By undercutting, it can capture the whole market and hence by choosing a small enough price reduction it can increase its profits.

Step 3. The competitive price is a Nash equilibrium. If both firms set the competitive price then neither can gain by raising its price. If one firm raises its price while the other continues to set $p_i = c$, the lower-priced firm will face the industry demand leaving the firm which has raised its price with no demand.

The Bertrand-Nash framework yields a very different relationship between structure and conduct from the Cournot-Nash equilibrium: with one firm the monopoly outcome occurs; with two or more firms the competitive outcome occurs. Large numbers are not necessary to obtain the competitive outcome and, in general, price-setting firms will set the competitive "price-taking" price.

Clearly it makes a difference whether firms choose prices or quantities. What grounds do we have for choosing between them? First, and perhaps most importantly, there is the question of the *type* of market. In some markets (for primary products, stocks and shares) the people who set prices (brokers) are different to the producers. There exists what is essentially an *auction* market: producers/suppliers release a certain quantity into the market and then brokers will sell this for the highest price possible (the market clearing price). The Cournot framework would thus seem natural where there are *auction* markets. While there are auction markets, there are also many industrial markets without "brokers" where the producers directly set the price at which they sell their produce. Clearly, the "typical" sort of market which concerns industrial economists is not an auction market but a market with price-setting firms. How can the use of the Cournot framework be justified in markets with price-setting firms?

It is often argued that the choice of Bertrand or Cournot competition rests on the *relative* flexibility of prices and output. In the Bertrand framework firms set prices and then produce to order. Thus, once set, prices are fixed while output is perfectly flexible. In the Cournot framework, however, once chosen, outputs are fixed while the price is flexible in the sense that it clears the market. Thus the choice between the two frameworks rests on the relative flexibility of price and output. This is of course an empirical question but many would argue that prices are more flexible than quantities (e.g. Hart 1985) and hence the Cournot equilibrium is more appropriate.

A very influential paper which explores this view is Kreps and Scheinkman (1983). They consider the subgame perfect equilibrium in a two-stage model. In the first stage, firms choose capacities; in the second stage firms compete with price as in the Bertrand model and can produce up to the capacity installed. The resultant subgame-perfect equilibrium of the two-stage model turns out to be equivalent to the standard Cournot outcome. This result, however, is not general and rests crucially on an assumption about contingent demand (the demand for a higher-priced firm given that the lower-priced firm does not completely satisfy its demand) – see Dixon (1987a). An alternative approach is to allow for the flexibility of production to be endogenous (Dixon 1985; Vives 1986). The Bertrand and Cournot equilibria then come out as limiting cases corresponding to when production is perfectly flexible (a horizontal marginal cost curve yields the Bertrand outcome) or totally inflexible (a vertical marginal cost curve at capacity yields Cournot).

Another reason that the Cournot framework is preferred to the Bertrand is purely technical: there is a fundamental problem of the non-existence of equilibrium in the Bertrand model (see Edgeworth 1925; Dixon 1987a). In our simple example firms have constant average/marginal costs. If this assumption is generalized – for example to allow for rising marginal cost – non-existence of equilibrium is a problem.[7]

A common argument for the Cournot framework is its "plausibility" relative to the Bertrand framework. Many economists believe that "numbers matter": it makes a difference whether there are two firms or two thousand. Thus the prediction of the Bertrand model – a zero price-cost margin with two or more firms – is implausible (see, for example, Hart 1979; Allen and Hellwigg 1986). The Cournot equilibrium captures the "intuition" that competition decreases with fewer firms. There are two points to be raised here: one empirical, one theoretical. Firstly, on the empirical level there exists little or no evidence that there is a smooth monotonic relationship between the level of concentration and the price-cost margin. Secondly, on the theoretical level, the stark contrast in the Bertrand and Cournot formulations has been exhibited here only in the case of a simple one-shot game. In a repeated game numbers may well matter. For example, Brock and Scheinkman (1985) consider a price-setting super-game and show that there is a relationship between numbers and prices that can be sustained in the industry (although the relationship is not a simple monotonic one). A related point is that the Nash equilibrium is a non-cooperative equilibrium. Numbers may well matter when it comes to maintaining and enforcing collusion and one of Bertrand's criticisms of Cournot was that collusion was a likely outcome with only two firms.

6.4 Cournot and Bertrand equilibria with differentiated commodities

In this section we will explore and contrast the Bertrand and Cournot approaches within a common framework of differentiated products with symmetric linear demands. As we shall see, there are again significant contrasts between markets where firms compete with prices and quantities. Firstly, we will compare the equilibrium prices and show that the Cournot equilibrium yields a higher price than the Bertrand equilibrium. Thus, as in the case of homogeneous products, Cournot competition is less competitive than Bertrand competition although the contrast is less.

Secondly, we contrast the "Stackelberg" equilibrium (where one firm moves before the other) and the corresponding "first-mover" advantage. In the Cournot framework the leader increases his own output and profits at the expense of the follower and total output increases, reflecting a more competitive outcome than the standard Nash equilibrium. In the Bertrand framework the Stackelberg leader will raise his price and increase his profits. The follower will also raise his prices and indeed his profits will increase by more than the leaders. Unlike the Cournot case there is then a "second-mover advantage" in the Bertrand case. Overall, with price competition the Stackelberg equilibrium leads to higher prices and profits and a contraction in total output. These differences between the behaviour of markets with price and quantity competition have important policy implications which will be discussed at the end of this section in the context of the recent literature on strategic trade policy.

We continue to assume that firms have constant average/marginal cost, A2. However, we will drop A1 and assume that there is a symmetric linear demand system; in the case of two firms with differentiated products we have:

*A*3 For 0 < a < 1

$$x_1 = 1 - p_1 + ap_2 \quad (11a)$$

$$x_2 = 1 - p_2 + ap_1 \quad (11b)$$

where a > 0 implies the two outputs are *substitutes* (e.g. margarine and butter): if a were negative then they would be complements (e.g. personal computers and software). In the exposition we will assume throughout that the firms produce *substitutes* and for technical reasons that a < 1 (i.e. quite plausibly the firm's own price has a greater absolute effect on its demand than the other firm's price).

The above equations express outputs (or more precisely, demands) as a function of prices.[8] If we want to explore the Cournot framework with differentiated products we need to invert (11) to give the prices that will "clear" the markets for chosen outputs.

Inverting (11) we have:

$$\begin{aligned} p_1 &= a_0 - a_1 x_1 - a_2 x_2 \\ p_2 &= a_0 - a_1 x_2 - a_2 x_1 \end{aligned} \quad (12)$$

where $a_0 = \frac{1+a}{1-a^2}$; $a_1 \frac{1}{1-a^2}$; $a_2 = \frac{a}{1-a^2}$.

Since a > 0 both prices are decreasing in both outputs. Thus an increase in x_1 by one unit will decrease p_1 by a_1, and p_2 by a_2 (of course $a_1 > a_2$ for a < 1).

6.4.1 Cournot-Nash equilibrium

There are two firms which choose outputs, the resultant prices given by the inverse demand system (12). Firm 1's "payoff" function is:

$$p_1 = x_1\left[\left(a_0 - a_1 x_1 - a_2 x_2\right) - c\right]$$

To obtain firm 1's reaction function x_1 is chosen optimally given x_2:

$$\frac{\partial p_1}{\partial x_1} = a_0 - 2a_1 x_1 - a_2 x_2 - c = 0$$

Solving for x_1 this yields:

$$x_1 = r_1(x_2) = \frac{a_0 - c - a_2 x_2}{2a_1} \tag{13}$$

The slope of the reaction function is given by:

$$\left.\frac{dx_1}{dx_2}\right|_{r_1} = \frac{-a_2}{2a_1} = -\frac{a}{2} < 0$$

With substitutes each firm's reaction function is downward sloping in output space as in Figure 2.

The firms are identical and there is a unique symmetric equilibrium at N with $x_1 = x_2 = x^c$:

$$x^c = \frac{a_0 - c}{2a_1 + a_2} = \frac{1 + a \ - c\left(1 - a^2\right)}{2 + a} \tag{14}$$

with resultant price:

$$p^c = \frac{1 + c(1-a)}{(2+a)(1-a)} \tag{15}$$

Figure 2 Cournot reaction functions

6.4.2 Bertrand-Nash equilibrium

Turning now to the Bertrand case firms choose *prices* so that we use the direct demand system (11). Firm 1's profits are:

$$\mathbf{p}_1 = p_1(1 - p_1 + ap_2) - c(1 - p_1 + ap_2) \tag{16}$$

$$\frac{\partial p_1}{\partial p_1} = 1 - 2p_1 + ap \ \ + c = 0$$

Hence firm 1's reaction function in price space is:

$$p_1 = s_1(p_2) = \frac{1 + c + ap_2}{2} \tag{17}$$

The slope is:

$$\left.\frac{dp_1}{dp_2}\right|_{s_1} = \frac{a}{2} > 0$$

Thus the two firms' reaction functions are upward sloping in price space as depicted in Figure 3.

Figure 3 Bertrand-Nash equilibrium

There is a unique symmetric equilibrium price $p_1 = p_2 = p^b$ with corresponding output, price and price-cost margins:

$$p^b = \frac{1+c}{2-a} \qquad (18)$$

$$x^b = \frac{1-c(1-a)}{2-a} \qquad (19)$$

$$m^b = \frac{1-c(2-a)}{1+c} \qquad (20)$$

How do the Cournot and Bertrand equilibria compare? Direct comparison of (18) with (15) shows that $p^b < p^c$: i.e. the Bertrand equilibrium prices are *lower* than Cournot prices. We formulate this in the following observation:

Observation. If firms' demands are interdependent a ? 0 then:

$$p^c > p^b; \quad x^c < x^b; \quad m^c > m^b$$

If a = 0 then each firm is, in effect, a monopolist since there are no cross-price effects and the two outcomes are, of course, the same. It should be noted that the above observation remains true when the goods are complements (-1 < a < 0).

With differentiated products, then, Bertrand competition will be more competitive than Cournot competition although the difference is less stark than in the case of homogeneous products. With product differentiation firms have some monopoly power even with price competition and do not have the same incentives for undercutting their competition as in the homogeneous goods case.

What is the intuition behind the above observation that price competition is more competitive than quantity competition? Clearly, for a monopoly, it makes no difference whether price or quantity is chosen; it simply chooses the profit-maximizing price-output point on its demand curve. There is a sense in which this is also true for the oligopolist: *given* what the other firm is doing it faces a demand curve and chooses a point on that demand curve. However, the demand curve facing firm 1 will be different if firm 2 keeps x_2 constant (and hence allows p_2 to vary) from when firm 2 keeps p_2 constant (and hence allows x_2 to vary). From (11) if firm 2 has price as its strategy and holds p_2 constant, firm 1's demand is:

$$x_1 = (1 + ap_2) - p_1 \tag{21}$$

with slope

$$\frac{dx_1}{dp_1}\bigg|_{p_2} = -1 \tag{22a}$$

and elasticity

$$h_1\big|_{p_2} = \frac{p_1}{x_1} \tag{22b}$$

If, on the contrary, firm 2 has output as its strategy it allows its price p_2 to vary as p_1 varies (to keep x_1 constant):

$$p_2 = 1 - x_2 + ap_1 \tag{23}$$

Substituting (23) into (11a) we obtain firm 1's demand when x_2 is held constant:

$$x_1 = (1 + a) - (1 - a^2)p_1 - ax_2 \tag{24}$$

with slope and elasticity:

$$\frac{dx_1}{dp_1}\bigg|_{x_2} = -(1 - a^2); \qquad h_1\big|_{x_2} = -\frac{p_1}{x_1}(1 - a^2) \tag{25}$$

Clearly, comparing elasticities (22) and (25):

$$h_1\big|_{p_2} < h_1\big|_{x_2} < 0$$

Thus the demand facing firm 1 is *more* elastic when firm 2 holds p_2 constant (and allows x_2 to vary) than when x_2 is held constant (and p_2 allowed to vary). For example, suppose that firm 1 considers moving up its demand curve to sell one less unit of x_1 with substitutes (a > 0). If firm 2 holds x_2 constant then as firm 1 reduces its output and raises its price the price for x_2 will rise (via (11b)). Clearly the demand for firm 1 will be more elastic in the case where firm 2 does not raise its price and expands output.

We have derived the above observation under very special assumptions A1, A3. How far can we generalise this comparison of Cournot and Bertrand prices? This has been the subject of much recent research – see for example Cheng (1984), Hathaway and Rickard (1979), Okuguchi (1987), Singh and Vives (1984), Vives (1985a,b). Vives (1985a) considers a more general differentiated demand system which need not be linear or symmetric (*ibid.* 168) and derives fairly general conditions for which the Bertrand price is less than the Cournot price. Of course there need not be unique Cournot or Bertrand equilibria: with multiple equilibria the comparison becomes conceptually more complex. Vives (1985b) has established a result that for very general conditions there exists a Bertrand equilibrium which involves a lower price than any Cournot equilibrium.

Of course there are other contrasts to be drawn between Cournot and Bertrand-Nash equilibria. For example, there is the question of welfare analysis employing standard consumer surplus. A simple example employing the linear demand system (11), (12) is provided by Singh and Vives (1984: 76) which shows that the sum of consumer and producer surplus is larger in Bertrand than in Cournot-Nash equilibrium both when goods are substitutes and complements.

6.4.3 Stackelberg leadership and the advantages of moving first

The difference between Cournot and Bertrand competition go deeper than the simple comparisons of the previous section. To illustrate this we will examine the advantages of moving first in the two frameworks. The standard Nash equilibrium assumes that firms move simultaneously. However, Heinrich von Stackelberg (1934) suggested an alternative in which one firm (the leader) moves first, the other (the follower) moves second. Thus when the follower chooses its strategy it treats the leader's choice as given. However, the leader will be able to infer the follower's choice and take this into account in its decision. The explicit algebraic analysis of the

Stackelberg equilibrium is rather complicated and we will rather employ the familiar iso-profit loci.[9] In the following analysis it is important to note that under A1, A3 the model is perfectly symmetric; thus whether in price or quantity space the firms' reactions functions are "symmetric" in the sense that firm 1's reaction is a reflection of firm 2's in the 45^o line (see Figures 2,3). Similarly, firm 1's iso-profit loci are simply reflections of firm 2's in the 45^o line and vice versa.

Firstly we analyse the Stackelberg equilibrium in the Cournot case. The follower (firm 2) will simply choose its output to maximize its profits given x_1 so that $x_2=r_2(x_1)$. The leader, however, will choose x_1 to maximize its profits given that x_2 depends on x_1 via r_2. Thus by moving first the leader can pick the point on firm 2's reaction function that yields it the highest profits: this is represented in Figure 4 by the tangency of iso-profit loci p^L to r_2 at point A.

Figure 4 First-mover advantage in Cournot model

If firm 2 were the leader and firm 1 the follower then, by symmetry, firm 2's Stackelberg point would be at point B – the reflection of A in the 45^o line (at this point firm 2's iso-profit locus is tangential to firm 1's reaction function). Comparing points A, B and the Nash equilibrium at point N we can see that if firm 1 is the leader it earns p^L which is greater than in the Nash equilibrium p^N. If firm 1 is a follower it will end up at point B and earn only p^F which is *less* than p^N. Hence, in the Cournot framework we have:

$$p^L \;>\; p^N \;>\; p^F \qquad (\textit{Cournot})$$

profits of leader	Nash profits	profits of follower

There is thus a first-move advantage in two senses: the leader earns more than in the simultaneous-move case ($p^L > p^N$); the leader earns more than the follower ($p^L > p^F$). The leader increases his output and profits at the expense of the follower (in fact, the decline in the follower's profits from p^N to p^F is *larger* than the increase in the leader's from p^N to p^L: industry profits fall).

In the Bertrand case the story is rather different: there is a "second-mover" advantage. The reaction functions and iso-profit loci of firm 1 depicted in price space are shown in Figure 5 and again are symmetric. The iso-profit loci for firm 1 are higher the further away they are positioned from the x-axis (firm 1 will earn higher profits the higher p_2 is). N is the Nash equilibrium, A occurs if 1 is the leader, B if 2 is the leader.

Figure 5 Second-mover advantage in Bertrand model

If firm 1 is the leader it will choose to raise its price moving along firm 2's reaction function S_2 from N to A and profits will increase from p^N to p^L. However, if firm 1 is the follower it will end up at B with profits p^F. Note that the follower raises his price by less than the leader and that $p^F > p^L$. thus the leader will set a higher price than the follower, produce a lower output and earn *less* profits:

$$p^F > p^L > p^N \qquad (\textit{Bertrand})$$

There is an advantage to moving "second". The second-mover advantage goes beyond Bertrand equilibria and extends to any game with upward-sloping reaction functions (Gal-Or) 1986). There is still a first-mover advantage in the sense that the leader earns more than in the simultaneous move case ($p^L > p^N$). In contrast to the Cournot case, in the Bertrand case Stackelberg leadership leads to higher prices, profits and lower outputs.

6.4.4 Prices vs. quantities

Price and quantity competition have very different implications for the nature of product market competition between firms. Most importantly, from the firms' point of view, price competition leads to lower profits than does quantity competition in Nash equilibrium. As was discussed earlier, whether firms should be viewed as competing with price or quantity can be seen as depending on structural or institutional characteristics of the market – the flexibility of production, whether the market is an auction market etc.

An alternative approach is to treat the firm's decision to choose price or quantity as itself a strategic decision (Klemperer and Meyer 1986; Singh and Vives 1984). While it is perhaps not quite clear how firms might achieve this, it is at least a useful "experiment" and will reveal the incentives which firms have to achieve one or the other type of competition.

Without uncertainty, this "experiment" is not fruitful: firms are *indifferent* between choosing price or quantity. The reason is that from the individual firm's perspective it simply faces a demand curve and, like a monopolist, chooses a point on that demand curve. It can achieve any point on the demand curve by choosing either price or quantity. The firm's own price/quantity decision does not affect this demand curve which is rather determined by the *other* firm's choice. Firm 1's choice has a pure externality effect on the demand faced by firm 2: if 1 chooses price, 2's demand is more elastic than if firm 1 had chosen quantity. However, in the Nash framework, each firm will ignore this externality: given the other firm's choice, each firm will face a particular demand curve and will be indifferent between setting price or quantity itself. In the case of duopoly there will be four Nash equilibria in this strategic game: one where both set quantities (Cournot); one where both set prices (Bertrand); and two asymmetric equilibria where one sets price and the other quantity. With certainty, then, allowing firms to choose price or quantity tells us nothing about which may be more appropriate.

The presence of uncertainty (adding a stochastic term to A3, for example) can mean that firms have a strict preference between price and quantity setting. The results depend very much on the exact assumptions made (is demand uncertainty additive or multiplicative; is demand linear in prices?). For the simple linear demand system A3 with an additive stochastic term firms will prefer quantity setting if marginal costs are increasing; they will be indifferent if marginal costs are constant (as in A1); they prefer price setting if marginal costs are decreasing (Klemperer and Meyer 1986; proposition 1). While this and related results are at present rather specific they do suggest that the presence and nature of uncertainty provide some insights into how firms view the alternatives of price and quantity setting.

6.5 Precommitment: strategic investment and delegation

In the previous section we explored the nature of the first-mover advantage in the Cournot and Bertrand framework. Clearly, if we start from the Nash equilibrium there is an incentive for the firm to precommit its output/price to obtain this first-mover advantage. By "precommitment" it is meant that the firm takes some action prior to competing in the product market which commits it to a certain course of action. In the standard Cournot model it is not credible for one firm to produce the

Stackelberg output in the simultaneous move game. For example, in terms of Figure 6, firm 1's Stackelberg point A is not on its reaction function – so that given firm 2's output x_{2A} firm 1 would like to produce x_1. The only credible equilibrium is the Nash equilibrium at N. In order to move towards its Stackelberg point the firm must be able to precommit its output in some way. In the previous section we simply assumed that the leader was able to move first. In some situations it is natural to assume a particular sequence of moves (e.g. entrant/incumbent, dominant firm). However, in the case of active incumbents which are competing on even terms, simultaneous moves seem more natural.

Figure 6 Non-credibility of Stackelberg point

Given that there is an incentive for the firm to precommit how can this be achieved? This section will look at two methods of precommitment which have received much recent attention – precommitment through investment and precommitment through delegation. The basic idea is simple: the firm can take actions prior to competing in the market which will alter the Nash equilibrium in the market. Firms can take actions such as investment decisions[10] and choice of managers that are irreversible (in the sense of being "fixed" over the market period) and which alter the firm's reaction function thus shifting the Nash equilibrium in the market. We will first consider how investment by firms can be used strategically to alter the market outcome.

For a wide range of industrial processes economists since Marshall have taken the view that it is appropriate to treat the capital stock decision as being taken on a different time scale (the "long run") to price/output decisions (the "short-run"). When firms compete in the product market it follows that they treat their capital stock as fixed. The capital stock chosen by the firm will influence its *costs* when it competes in the market. The fact that capital is committed "before" output/price decisions means that it can use investment strategically to influence the market outcome. In essence, through its choice of capital stock, the firm will determine the short-run costs which it will have when it chooses output/price; the firm's marginal costs will determine its reaction function and hence the Nash equilibrium in the product market. Schematically:

Investment ? short-run marginal cost ? reaction function ? market equilibrium

For example, in the Cournot case the firm can increase its investment, reduce its marginal cost and hence shift its reaction function out so that the product market equilibrium moves towards the Stackelberg point. Of course, this precommitment is not costless: capital costs money and, as we shall see, such use of capital leads to productive inefficiency. Again there is an important dichotomy between the Cournot and Bertrand approaches: if the product market is Cournot then the firm will want to overinvest; if the product market is Bertrand then the firms will want to underinvest. We will briefly illustrate both situations.

The structure of strategic investment models is very simple: there are two stages to capture the distinction between the short and the long run. In the first "strategic" stage, the firms choose their capital stock; in the second "market" stage firms choose output/price. The choice of capital stock in the first stage will determine the cost function which the firm has. In the previous two sections we assumed constant average/marginal cost at c (A2). This can be conceived of as the long-run cost function. In order to keep the exposition consistent we will assume that firms have a production function of the form:

*A*4 $\quad x_1 = k_i^{1/2} L_i^{1/2}$

$$c(x_i, k_i) = rk_i + \frac{x_i^2}{k_i} \tag{26}$$

where linear increasing marginal cost:

$$\frac{\partial c}{\partial x_i} = \frac{2}{k_i} x_i \tag{27}$$

Thus an increase in investment lowers the marginal cost of producing output. The production function A4 displays constant returns to scale and hence the long-run cost function has constant average/marginal cost in terms of A2,[11] minimum average cost $c = 2\text{v}r$.

We will first outline the strategic investment model with Cournot competition in the product market, a simple version of Brander and Spencer's (1983) article. If investment is used non-strategically, then the firm simply operates on its long run cost function given by A2: capital and labour are chosen to minimise production costs. In the strategic investment framework, however, the firm's costs will be given by its short-run cost function (26). Turning to the market stage, the firm's profits are:

$$\pi_i = x_i\left(a_0 - a_1 x_i - a_2 x_j\right) - rk_i - \frac{x_i^2}{k_i} \qquad (28)$$

Figure 7 Investment shifts from 1's reaction function out

The reaction function which the firm has in the market stage, conditional on k_i, is derived by setting $\partial \pi_i / \partial x_i = 0$:

$$x_i = r_i\left(x_j, k_i\right) = \frac{a_0 - a_2 x_j}{2a_1 + \left(2/k_i\right)} \qquad (29)$$

By increasing its investment firm *i* will reduce its marginal costs and from (29) it will shift its reaction function out as in Figure 7. Given the level of investment by the two firms (k_1 k_2) are obtained by solving (29) for x_1 and x_2 (we leave this as an exercise for the reader). In general form:

$$x_i = x_i\left(\underset{+}{k_i}, \underset{-}{k_j}\right) \qquad (30)$$

This notation signifies that firm *i*'s equilibrium output in the market stage depends positively on its own investment and negatively on investment by the other firm. Suppose we start off at point A in Figure 9: an increase in k_1 to $k_1^{'}$ shift out r_1 so that the market equilibrium goes from A to B, x_i rising and x_2 falling. Conversely, an increase in k_2 shifts the equilibrium from A to C. Thus by altering their investment, the firms can alter their reaction functions and hence the market stage equilibrium.

Figure 8 Market stage equilibrium given investment k_1, k_2

How is the optimal level of investment in the first strategic stage determined? Firms choose investment levels k_i *given* that the second-stage outputs will be as in (30). We can see their profits as a function of capital stocks. For firm 1 we have profits:

$$p_1(k_1,k_2) = x_1(k_1,k_2)[a_0 - a_1x_1(k_1,k_2) - a_2x_2(k_1,k_2)] - c(x_1(k_1,k_2),k_1) \quad (31)$$

The RHS term in square brackets is the price, which is multiplied by output to obtain revenue from which are subtracted costs.

The firm will choose k_1 to maximize its profits (31) hence:

$$\frac{\partial p_1}{\partial k_1} = \frac{\partial x_1}{\partial k_1}\left[a_0 - 2a_1x_1 - a_2x_2 - \frac{\partial c}{\partial x_1}\right] - a_2x_1\frac{\partial x_2}{\partial k_1} - \frac{\partial c}{\partial k_1} = 0 \quad (32)$$

Since firm 1 chooses x_1 to maximize profits given x_2 and k_1 in the second stage, the bracket on the right-hand side of (32) is zero (it is simply its reaction function (29)). Hence (32) becomes:

$$\frac{\partial c}{\partial k_1} = -a_2\frac{\partial x_2}{\partial k_1}x_1 > 0 \quad (33)$$

What does (33) tell us? $\partial c/\partial k_1$ gives the effect of investment on the total costs of producing x_1. If $\partial c/\partial k_1 = 0$, as in the standard non-strategic case, then k_1 *minimizes* the cost of producing x_1. If $\partial c/\partial k_1 > 0$, as in (33), then there is "overcapitalization": more investment than would minimize costs (a reduction in k_1 would reduce average costs). If $\partial c/\partial k_1 < 0$, then there is "undercapitalization": less investment than would minimize costs.

Figure 9 Market stage equilibrium and changes in investment

With a Cournot market stage, then, there is overcapitalization of production in the market stage. The intuitive reason is quite simple. Given the other firm's reaction function, each firm can shift its own reaction function out towards its Stackelberg point. Of course there is a cost to: more investment leads to higher capital costs and inefficient production. The firms will shift out their reaction functions beyond their "innocent" level and the final product market equilibrium will be at a point such as S in Figure 10. At the equilibrium level of investment the additional cost of investment equals the additional gains from moving out the reaction function further. In the strategic investment equilibrium S then, both firms produce a larger output than in the non-strategic equilibrium N.

In the Bertrand case an exactly analogous argument applies for strategic investment. However, there is the opposite result of *undercapitalization*. The Stackelberg equilibrium results in higher prices and lower outputs than in the Bertrand case. Thus firms will *restrict* investment relative to the innocent Bertrand equilibrium in order to shift their reaction function *out* in price space, as in Figure 11. Starting from the “innocent” Bertrand equilibrium at N, if firm 1 restricts its investment its marginal costs *rise* and its reaction function shifts outwards to $s_1^{'}$ (an outward shift because with higher marginal costs it will wish to set a higher price and produce a smaller quantity given the price chosen by the other firm). If both firms underinvest strategically the resultant equilibrium will be at S with higher prices and lower output.

Figure 10 Strategic investment equilibrium – the Cournot case

Figure 11 Underinvestment raises prices

We will briefly sketch the algebra underlying this result. Under A1 and A4 the firm’s profits are:

$$\mathbf{p}_i = p_i - p_i^2 - a p_i p_j - \frac{1}{k_i}\left(1 - p_i - a p_i p_j\right)^2 - r k_i \qquad (34)$$

Setting $\partial \mathbf{p}_i / \partial p_i = 0$ firm i’s reaction function $p_i = s_i(p_j)$ is:

$$P_i = \frac{1+(2/k_i)}{2+(2/k_i)} + \frac{a\left(1+(2/k_i)\right)}{\left(2+(2/k_i)\right)} P_j \qquad (35)$$

Solving for p_i to p_j given k_i and k_j in general terms we have:

$$P_i = p_i\left(\underset{-}{k_i}, \underset{+}{k_j}\right) \qquad i, j = 1, 2, i ? j \qquad (36)$$

The firms choose k_i to maximize (34) given that in the market subgame firms’ prices are given by (36) (i.e. a Bertrand-Nash equilibrium occurs).

$$\frac{d\mathbf{p}}{dk_i} = \frac{\partial \mathbf{p}_i}{\partial p_i} \cdot \frac{dp_j}{dk_i} + \frac{\partial \mathbf{p}_i}{\partial p_j} \cdot \frac{dp_j}{dk_i} + \frac{\partial \mathbf{p}_i}{\partial k_i} = 0 \qquad (37)$$

(0) (-) (+) (-) (+)

Note that $\partial \mathbf{p}_i / \partial p_i = 0,$ since firms are in their market stage reaction function (35) and, further, that $\partial \mathbf{p}_i / \partial k_i = -\partial c_i / \partial k_i$. Hence (37) can be expressed as:

$$\frac{\partial c}{\partial k_i} = \frac{\partial \mathbf{p}_i}{\partial p_j} \frac{dp_j}{dk_i} < 0 \qquad (36)$$

(+) (-)

That is, *undercapitalization* of production results from the strategic use of investment with Bertrand product market competition.

Clearly, the result of strategic investment models depends on the nature of product market competition. Other papers have made different assumptions than the simple Cournot-Nash and Bertrand-Nash equilibria. Dixon (1985) considers the case of a *competitive* product market; Eaton and Grossman (1984) and Yarrow (1985) a *conjectural* Cournot equilibrium; Dixon (1986b) a *consistent conjectural* variation equilibrium in the product market.

Since production will generally be inefficient in a strategic investment equilibrium, firms have an incentive to try and precommit their labour input at the same time as their capital. By so doing, firms will be able to produce any output efficiently, while being free to precommit themselves to a wide range of outputs. In Dixon (1986a), precommitment is treated as a strategic choice by the firm: the firm can precommit either, neither, or both capital and labour in the strategic stage. Because of the strategic inefficiency in production that occurs when only capital is precommited, under almost any assumption about the nature of product market competition, firms would prefer to precommit both factors of production (Dixon 1986a; Theorem and ? p. 67). If firms precommit both factors of production in the strategic stage then, in effect, they have chosen their output for the market stage and the resultant equilibrium is equivalent to the standard Cournot equilibrium. How might firms be able to precommit their output in this manner? One important method that may be available is choice of technology. More specifically, the firm may have a choice between a putty-putty technology that allows for smooth substitution of capital for labour in the market stage or an otherwise equivalent putty-clay technology that is Leontief in the market stage. If the firm chooses a putty-clay technology then its choice of investment and technique in the strategic stage effectively ties down its output and employment in the market stage. If possible, then, firms would prefer to have totally inflexible production in the market stage. This strong result ignores uncertainty of course. If demand or factor prices are uncertain there will be a countervailing incentive to retain flexibility.

In strategic investment models it is firms themselves which precommit. Governments, however, can undertake precommitments which firms themselves cannot make. In the context of trade policy there has been much recent research on how governments can improve the position of their own firms competing in international markets (see Venables 1985 for excellent surveys). If domestic firms are competing in foreign markets the net benefit to the home country in terms of consumer surplus is the repatriated profits – total revenue less the production costs (with competitive factors markets production costs represent a real social cost to the exporting country). Government trade policy may therefore be motivated by what is called “rent extraction”: that is, helping their own firms to make larger profits which are then repatriated. Trade policy, usually in the form of an export subsidy or tax, is a form of precommitment by the government which enables domestic firms to improve their position in foreign markets. Brander and Spencer (1984) presented the first model based on the rent-extraction principle and argued for the use of export subsidies in the context of a Cournot-Nash product market. Subsidies have the effect of *reducing* the marginal costs faced by exporters and can thus be used to shift out their reaction functions to the Stackelberg point (the cost subsidies “cost” nothing from the point of view of the exporting country since they merely redistribute money from the taxpayers to shareholders). As Eaton and Grossman (1983) argued, the exact form of the trade policy will be sensitive to the nature of product market competition. With a Bertrand product market, of course, rent extraction arguments lead to the imposition of an export *tax* since this will shift the Bertrand competitor’s reaction function outwards in price space towards its Stackelberg point.

The incentive to precommit in oligopolistic markets also sheds light on one of the perennial issues of industrial economics – what are the objectives of firms? The divorce of ownership from control can be viewed as an act of delegation by shareholders. This act of delegation can be employed as a form of precommitment by shareholders. What sort of Managers do shareholders want to manage their firms? There is an obvious answer to this question which underlies the managerialist view of Marris (1964): shareholders want managers who maximize profits (share valuation) and work hard. This may be true in the context of monopoly: in an imperfectly competitive framework matters are rather different. Several recent papers (Fershtman 1985; Lyons 1986; Vickers 1985a) have shown how higher profits for shareholders can be obtained when they have non-profit maximizing shareholders. The reaction functions of firms in the standard Cournot and Bertrand models are based on the assumption of profit maximization. By choosing managers with different objectives (e.g. a preference for sales, or an aversion to work) the firms' reaction functions will be shifted. We will illustrate this with a very simple example adapted from Lyons (1986). Managers maximize utility which depends on profits (remuneration) and sales R (power, prestige and so on). The utility is a convex combination of the two:

$$\begin{aligned} u &= yp + (1-y)R \qquad 0 \le y \le 1 \\ &= R - yc \end{aligned}$$

since $p = R - c$. The coefficient y represents the weight put on *profits*: $y = 1$ is profit maximization, $y = 0$ yields sales maximization.

Figure 12 Equilibrium outcome and managerial preferences

Using the common framework A1, A2, assuming that managers choose outputs to maximize utility, we can derive the firm's reaction functions:

$$u_1 = x_1(a_0 - a_1 x_1 - a_2 x_2) - ycx_1$$

$$\frac{\partial u_1}{\partial x_1} = a_0 - 2a_1 x_1 - a_2 x_2 - yc = 0$$

which yields the reaction function:

$$x_1 = r_1(x_2, y_1) = \frac{a_0 - yc - a_2 x_2}{2a_1}$$

(-) (-)

By choosing managers with a preference for sales (i.e. y smaller than unity) shareholders can push out their firm's reaction function. In Figure 12 we depict the two extreme reaction functions: the one nearest the origin corresponding to profit maximization $y = 1$, the other to sales maximization $y = 0$. Given firm 2's reaction function, firm 1 can move to any point between N and T by choosing the appropriate value of y. If, as depicted, the Stackelberg point A lies between N and T, then firms will be able to attain their Stackelberg point – note that since A will lie to the left of N, the choice of y will surely be less than unity, reflecting non-maximization of profits due to some sales preference. In such a market, if one firm is a profit maximizer with $y = 1$ and the other has management with $y < 1$ the non-profit maximizing firm will earn *more* than the profit-maximizing firm! Of course, in a Bertrand market the shareholders would wish to choose managers who would restrict output and raise prices – perhaps lazy managers with an aversion to work (see Dixon and Manning 1986, for an example). While we have talked about different "types" of managers, the precommitment made by shareholders can be seen as taking the form of different types of remuneration packages which elicit the desired behaviour from managers.

In an imperfectly competitive market then, it can pay shareholders to have non-profit maximizing managers. There need not be the conflict of interest between owners and managers that is central to managerialist theories of the firm. Also, "natural selection" processes need not favour profit maximizers in oligopolistic markets since, for example, sales-orientated managers can earn larger profits than their more profit-orientated competitors. This is a comforting result given the apparent prevalence of motives other than profits in managerial decisions.

The presence of a first-mover advantage means that firms competing in an oligopolistic environment have an incentive to precommit themselves in some way. We have explored *two* methods of precommitment: through investment and through delegation. Strategic investment leads to productive inefficiency and, from the point of view of the firm, it may be cheaper to make its precommitment through its choice of managers rather than its choice of capital stock.

6.6 Competition over time

In general, Nash equilibria are "inefficient" in the sense that in equilibrium profits of all the firms can be increased. The fundamental reason is that firms' profits are interdependent (via the payoff function): each firm's profits depend partly on what the other firms are doing. There is thus an "externality" involved when each firm chooses its strategy. For example, in the Cournot framework if one firm raises its output, it reduces the prices obtained by the other firms, thus reducing their profits (a negative externality). In the Bertrand case, a rise in price by one firm is a positive externality since it raises the demand for other firms. Under the Nash assumption, each firm chooses its own strategy taking into account only the impact on its *own* profits, ignoring the externality.

The inefficiency of Nash equilibria can easily be demonstrated using the abstract notation of the section "Non-cooperative equilibrium". For simplicity we will take the case of duopoly. To obtain an *efficient* (Pareto optimal) outcome between the two firms, simply maximize a weighted sum of firms' profits:

$$\max_{a_1, a_2} \; l p_1(a_1, a_2) + (1 - l) p_2(a_1, a_2) \tag{39}$$

where $0 \leq l \leq 1$. The first-order conditions for (39) are:

$$l \frac{\partial p_1}{\partial a_1} + (1 - l)\frac{\partial p_2}{\partial a_1} = 0 \tag{40a}$$

$$l \frac{\partial p_1}{\partial a_2} + (1 - l)\frac{\partial p_2}{\partial a_2} = 0 \tag{40b}$$

Figure 13 The profit frontier

The leading diagonal terms represent the effect of a_i on p_i, the firm's strategy on its own profits. The off-diagonal terms reflect the "externality", the effect of a firm's strategy on the other firm's profits. Depending on the weight l, a whole range of Pareto-optimal outcomes is possible (corresponding to the contract curve of the Edgeworth box). These outcomes can be represented as the profit frontier in payoff space, as in Figure 13. On the frontier, each firm's profits are maximized given the other firm's profits. As l moves from 0 to 1, more weight is put on firm 1's profits and we move down the profit frontier.[12]

The Nash equilibrium profits are not Pareto optimal and lay *inside* the profit frontier at point N, for example. To see why note that for a Nash equilibrium to occur, both firms choose a_i to maximize their own profits (they are both on their reaction functions). Thus the first-order equations defining the Nash equilibrium are:

$$\frac{\partial p_1}{\partial a_1} = \frac{\partial p_2}{\partial p_2} = 0 \qquad (41)$$

If we compare (41) with (40) we can immediately see that if there is some interdependence captured by a non-zero cross-effect $\left(\partial p_i / \partial a_j \neq 0\right)$ then (41) will not be efficient. If there is a *negative* cross-effect, then at the Nash equilibrium N:

$$\underset{(0)}{\frac{l \partial p_1}{\partial a_1}} \Big| N + (1-l) \underset{(-)}{\frac{\partial p_2}{\partial a_1}} \Big| N \langle 0 \qquad (42)$$

The marginal effect of a_i on the weighted sum of industry profits is *negative*: there is too much output chosen. Conversely, in the Bertrand case at the Nash equilibrium, the marginal effect of a price rise on the weighted sum of industry profits is *positive.*

This inefficiency of Nash equilibria means that there is an incentive for firms to collude – to choose their strategies (a_1, a_2) *jointly* and move from N towards the profit frontier. Of course, if the two firms could merge or write legally binding contracts, it would be possible for them to do this directly. However, anti-trust law prevents them from doing so: firms have to behave non-cooperatively. However, since the efficient outcomes are not Nash equilibria, each firm will have an incentive to *deviate* from the efficient outcome: it will be able to increase its profits, for example from (40), at an efficient outcome where partial $\partial p_i / \partial a_j$ is positive (negative) then $\partial p_i / \partial a_i$ will be negative (positive) so that a slight reduction (increase) in a_i will increase *i*' profits.

Given that firms have an incentive to cooperate, how can they enforce cooperative behaviour if there is also an incentive for firms to deviate from it? One response is to argue that firms compete over time: firms can enforce cooperative behaviour by punishing deviation from a collusive outcome. Since firms are involved in a repeated game, if one firm deviates at time *t*, then it can be "punished" at subsequent periods.

In a repeated game, might not such "threats" enable firms to enforce a collusive outcome over time? This question has provided the impetus for much research in recent years.

In a finitely repeated game with perfect information it turns out that the unique subgame-perfect equilibrium will be to have the Nash equilibrium in each period (assuming that there is a unique Nash equilibrium in the constituent game). That is, if we restrict firms to *credible* punishment threats, then those credible threats will not enable the firms to do better than their Nash equilibrium profits in each period. the argument is a standard backwards induction argument. Consider the subgame consisting of the last period. There is a unique Nash equilibrium for this subgame which is that the firms play their Nash strategies. Any other strategy in the last period would not be "credible", would not involve all firms adopting their best response to each other. Consider the subgame consisting of the last two periods. Firms know that whatever they do in the penultimate period, the standard Nash equilibrium will occur next period. Therefore, they will want to choose their action to maximize their profits in the penultimate period given what the other firms do. If all firms do this the standard Nash equilibrium will occur in the penultimate period. By similar arguments as we go backwards, for any period t, given that in subsequent periods the Nash equilibrium will occur, the Nash equilibrium will occur in period t as well. Hence finite repetition of the game yields the standard Nash outcome in each stage of the history of the market. This backwards induction argument goes back to Luce and Raiffa's analysis of the repeated prisoner's dilemma (1957).

In finitely repeated games, then, there is no scope for threats/punishments to move firms' profits above their Nash level. The argument relied upon a known terminal period, "the end of the world". An alternative approach is to analyse *infinitely* repeated games, reflecting the view that market competition is interminable. This raises a different problem: there are generally many subgame perfect equilibria in infinitely repeated games. Clearly, the above backwards induction argument cannot be employed in infinitely repeated games because there is no last period to start from! It has proven mathematically quite complex to characterize the set of subgame perfect equilibria in infinitely repeated games. There are two types of results (commonly called "Folk theorems") corresponding to two different views of how to evaluate the firm's payoffs over an infinitely repeated game. One approach is to view the firm maximizing its discounted profits for the rest of the game at each period (Lockwood

1984; Abreu 1985; Radner 1986). The other is to view that the firm does not discount but maximizes its average per-period payoff.

Let us first look at the "Folk theorem" for infinitely repeated games without discounting which is based on work by Rubenstein (1979). The key reference point is the "*security level*" of firms: this represents the worst punishment that can be inflicted on them in the one-shot constituent game. This is the "minimax" payoff of the firm, the worst payoff that can be imposed on the firm given that it responds optimally to the other firm(s). For example, take the simple Cournot model: the lowest level to which firm 1 can drive firm 2 is zero – this corresponds to where firm 2's reaction function cuts the *x*-axis and firm 2's output and profits are driven to zero. In the framework we have employed each firms' security level corresponding to the worst possible punishment it can receive is equal to zero. An individually rational payoff in the constituent game is defined as a payoff which yields both firms their security level. The basic result is that *any* individually rational payoff in the constituent game can be "sustained" as a perfect equilibrium in an infinitely repeated game without discounting. By "sustained" it is meant that there corresponds equilibrium strategies that yield those payoffs for each firm. In our example this means that *any* combination of non-negative profits is possible! This will include the outcomes on the profit frontier of course but also outcomes that are far worse than the standard Cournot-Nash equilibrium! In terms of Figure 13, the whole of the area between the axes and the profit frontier (inclusive) represents possible payoffs of some subgame perfect equilibrium!

With discounting, the range of possible equilibria depends on the discount rate d. At period *t* the future discounted profits for the rest of the game are:

$$\sum_{s=0}^{\infty} d^{s} p_{it+s}$$

where 0=d<1 (if the interest rate is *r* then d = 1/(1+ *r*)). The larger is d, the more weight is put on the future: as d tends to one, we reach the no-discounting case (since equal weight is put on profits in each period); with d equal to zero, the future is very heavily discounted and the firm concentrates only on the current period. The analysis of infinitely repeated games with discounting is rather more complex than the no-discounting case, not least because it is more difficult to define the firm's security level which itself varies with d (see Fudenberg and Maskin 1986). The basic Folk theorem is that: (a) as d? 0, then the set of perfect equilibrium payoffs shrink to the one-shot Nash payoffs; (b) as d? 1, then any individually rational payoff is an equilibrium payoff. Again, the analysis is rather complicated here and the reader is referred to Lockwood (1987) for an excellent analysis of the issues. The basic message for games with discounting is that the set of perfect equilibria depends on the discount rate and may be very large.

From the point of view of industrial economics the game-theoretic results for repeated games are far from satisfactory. With finite repetition the equilibrium is the same as in the one-shot case: with infinite repetition there are far too many equilibria – almost anything goes! There seems to be little middle ground.

However, recent advances involving games of imperfect information may provide some answer to this dilemma (Kreps *et al.* 1982). The basic idea is very simple. Suppose that the firms are uncertain about each other's objectives. In a repeated game, firms can learn about each other's "character" from observing their actions through time. In this circumstance, firms are able to build up reputations. Let us take a very simple example: there are two firms A and B with two strategies, cooperate (*c*) or defect (*d*). The resultant profits of the two firms are of the familiar prisoner's dilemma structure:

		B	
		c	d
A	c	(1,1)	(-1,2)
	d	(2,-1)	(0,0)

Defection is the "dominant" strategy: whatever the other firm does, defection yields the highest profits hence the unique Nash equilibrium is for both firms to defect. This outcome is Pareto-dominated by the outcome where both firms cooperate. If there is perfect information and the game is repeated over time, then the unique subgame-perfect equilibrium is for both firms to defect throughout (by the standard backwards induction argument).

Now following Kreps *et al.* (1982) introduce some uncertainty. We will take the case where firms are uncertain about each other's motivation. In general, firms are of two types: a proportion a are "Rats" and play rationally; proportion (1 – d) are "Triggers" and play "trigger" strategies. A trigger strategy means that the firm will play cooperatively until the other firm defects, after which it will punish the defector by playing non-cooperatively for the rest of the game.

In a multi-period game like this where there is imperfect information, each firm may be able to infer the other firm's type from its past actions. For example, if one firm defects when they have previously both been cooperative, then the other firm can infer that the other firm is a Rat (since a Trigger only defects in response to an earlier defection). By playing cooperatively, then, a Rat can leave the other firm guessing as to his true type; if a Rat defects he knows he will lose his reputation and "reveal" his true nature.

To illustrate this as simply as possible, we will consider what happens when the above game is repeated for three periods and firms have discount rates d. For certain values of a and d it will be an equilibrium for both firms to cooperate for the first two periods and defect in the last period. Consider the following strategy from a Rat's point of view (a Trigger will of course follow a trigger strategy).

Period 1: Cooperate

Period 2: Cooperate if the other firm cooperated in period 1, defect otherwise.

Period 3: Defect

We will now show that this can be a perfect-equilibrium strategy for a Rat. Recall that the Rat does not know whether his opponent is a Trigger or a Rat following the same strategy.

In period 3 it is clearly subgame perfect to defect – whatever the type of the opponent, be he Rat or Trigger, defection is the dominant strategy and yields the highest payoff. In period 2 the decision is a little more complex. If the other firm (B, say) defected in period 1 then, of course, he has revealed himself to be a Rat and so defection is the best response for A for period 2. If firm B did not defect in period 1, then he may be a Trigger or a Rat (with probability (1 – a) and a respectively). If firm A defects in period 2 then whatever the type of firm B, it will earn two units in period 2 and nothing in period 3 (since firm B will retaliate whether a Rat or a Trigger). Its expected discounted profits are 2. If, however, firm A cooperates in period 2 it will earn 1 unit of profit in that period: in period 3 its profit will depend on firm A's type – with probability a the other firm is a Rat and will defect anyway: with probability (1 – a) the other firm is a Trigger and will cooperate in the last period. Thus if A cooperates in period 2 its expected period 3 profits are a0 + (1 – a)2. In period 2 firm A's expected discounted profits, if it cooperates, will be 1 + d(1 – a)2. Clearly, firm A will cooperate in period 2 if the expected discounted profits doing so exceed those from defection, i.e.

defect in period 2	$2 < 1 + d(1-a)2$	cooperate in period 2

This is satisfied for d(1 – a) > ½. In period 1 the decision is similar. If it defects in period 1 it earns 2 then nothing thereafter. If it cooperates, then it expects to earn 1 in period 1, 1 in period 2 (from the foregoing argument) and (1 – a)2 in period 3. The expected discounted profits from cooperation in period 1 are thus 1+ d+ d(1 – a)2. If d(1 – a) > ½ then, again, cooperation is period 1 yields higher expected profits than defection. Thus the above strategy is subgame perfect if the proportion of Triggers is high enough (1 – a) > ½ d.

With uncertainty then, it can be an equilibrium to have both firms cooperating initially during the game and only to defect towards the end of the game (the last period in the above example). The intuition is simple enough: by playing cooperatively in the first two periods the Rat hides his true nature from his competitors. There is a "pooling" equilibrium early on: both Rats and Triggers cooperate so that cooperation yields no additional information about the firm's type to alter the "priors" based on population proportions a, (1 – a). One problem with this account – for neo-classical economists at least – is the need to assume the existence of non-rational players to sustain the collusive outcome. This is a problem in two senses. Firstly, there are an indefinite number of ways to be non-rational: alongside the Trigger, the bestiary of the non-rational includes the "Tit-for-Tats" (Kreps *et al.* 1982) and many other fantastical possibilities. Secondly, the methodology of most economics is based on an axiom that all agents are rational maximizers. It might be said that all that is required for such equilibria is the *belief* that there are some non-rational players. While this may be so, it would seem less than satisfactory if the belief were not justified by the existence of the required proportion a of Triggers.

This sort of equilibrium with imperfect information is called a *sequential equilibrium* and has the added ingredient that firms use the history of the game to learn about each other's type by Bayesian updating. The equilibrium strategies in the example need not be unique: for some values of d and a it is also an equilibrium for Rats to defect throughout the game, as in the full-information case. However, there exists the possibility of sustaining cooperative behaviour for some part of the game even with a limited period of play. The use of sequential equilibria has been applied to several areas of interest and industrial economists – most notably entry deterrence (Milgrom and Roberts 1982a, b).

6.7 Conclusions

This chapter has tried to present some of the basic results in the recent literature on oligopoly theory in relation to product market competition. Given the vastness of the oligopoly literature past and present, the coverage has been limited. For those interested in a more formal game-theoretic approach, Lockwood (1987) is excellent (particularly on repeated games and optimal punishment strategies). On the growing literature on product differentiation, Ireland (1986) is comprehensive. Vickers (1985b) provides an excellent survey of the new industrial economics with particular emphasis on its policy implications.

ACKNOWLEDGEMENTS

The material of this chapter is based on MSc lectures given at Birkbeck over the period 1985-86. I would like to thank students for their comments and reactions from which I learnt a lot. I would also like to thank Ben Lockwood for invaluable comments as well as Bruce Lyons, Steve Davies and the editors. Errors, alas, remain my own.

[1] a^*_{-i} is the n-1 vector of all firms' strategies excepting i's.

[2] See Debreu (1952), Glicksberg (1952), and Fan (1952). Strict concavity is a stronger condition than we need – it can be relaxed to quasiconcavity.

[3] A sufficient condition for uniqueness is that each firm's reaction function is a "contraction mapping" – see Friedman (1978) for a formal definition.

4 Note the change in the use of the word "strategy". In a one-shot game, the firm's strategy is simply the action it pursues. In a repeated game "action" and "strategy" cease to be equivalent, "strategy" being its "game plan", the rule by which the firm chooses its action in each period.

[5] Bayesian updating means that firms have subjective probabilities which they update according to Bayes rule. Firms start the game with "prior" beliefs, and revise these to take into account what happens. This is a common way to model learning in neoclassical models.

[6] This simply states that x_i is chosen to equate marginal revenue with marginal cost.

[7] The reason for this non-existence is quite simple – step 3 of our intuitive proof breaks down and the competitive price need not be an equilibrium. The competitive price is an equilibrium with constant returns because when one firm raises its price the other is willing and able to expand its output to meet all demand. However, if firms have rising marginal cost curves they are supplying as much as they want to at the competitive price (they are on their supply functions). If one firm raises its price there will be excess demand for the firm(s) still setting the competitive price. The firm raising its price will thus face this unsatisfied residual demand and, in general, will be able to raise its profits by so doing (see Dixon 1987a: Theorem 1). One response to this non-existence problem is to allow for *mixed* strategies (rather than firms setting a particular price with probability one, they can set a range of prices each with a particular probability). Mixed-strategy equilibria exist under very general assumptions indeed (Dasgupta and Maskin 1986a,b) and certainly exist under a wide range of assumptions in the Bertrand framework (Dasgupta and Maskin 1986a; Dixon 1984; Dixon and Maskin 1985; Maskin 1986). However, the analysis of mixed-strategy equilibria is relatively complex and it has yet to be seen how useful it really is. It can be argued that it is difficult to see that mixed strategies reflect a genuine aspect of corporate policy.

[8] The standard models of Bertrand competition assume that outputs are demand-determined (see 11) : each firm's output is equal to the demand for it. This was the assumption made by Chamberlin (1933) in his analysis of monopolistic competition. This is appropriate with constant costs since firms will be willing to supply any quantity at the price they have set (for $p_i = c$, profits are increasing in output). More generally, however, it is very strong. Surely firms will only meet demand insofar as it raises the firm's profits. With rising marginal cost the output that the firm wishes to produce given the price it has set is given by its supply function (the output that the firm wishes to produce given the price it has set is given by its supply function). If demand exceeds this quantity and there is *voluntary trading* then the firm will turn customers away (otherwise marginal cost would exceed price). This approach is similar to Edgeworth's (1925) analysis of the homogeneous case – see Dixon (1987b), Benassy (1986). Benassy (1986) has analysed the implications of including an Edgeworthian voluntary trading constraint on price-setting equilibria. While the Nash equilibrium prices will be the same there is, however, an existence problem: if demand is highly cross-elastic between firms then no equilibrium may exist.

[9] The formula can be obtained by total differentiation of the implicit function $p_1(a_i, a_j) = q$.

[10] "Investment" can be taken as any fixed factor – (capital, R&D, firm-specific human capital and so on).

[11] Long-run average cost is derived as follows. Minimise total costs $rK + L$ with respect to the production function constraint A4. Since the production function displays constant returns, long- run average and marginal cost are equal.

[12] The profit frontier in Figure 13 is derived under the common framework A1-2. Linearity comes from constant returns with a homogeneous product. The actual solution is that *total* output on the frontier equals the monopoly output M, with total profits at their monopoly level µ. L determines firms' share of output and profit:

$$x_1 = lM; p_1 = lm; x_2 = (1-l)M; p_2 = (1-l)m$$

With diminishing returns, i.e. a strictly convex function, the profit frontier will have a concave shape.

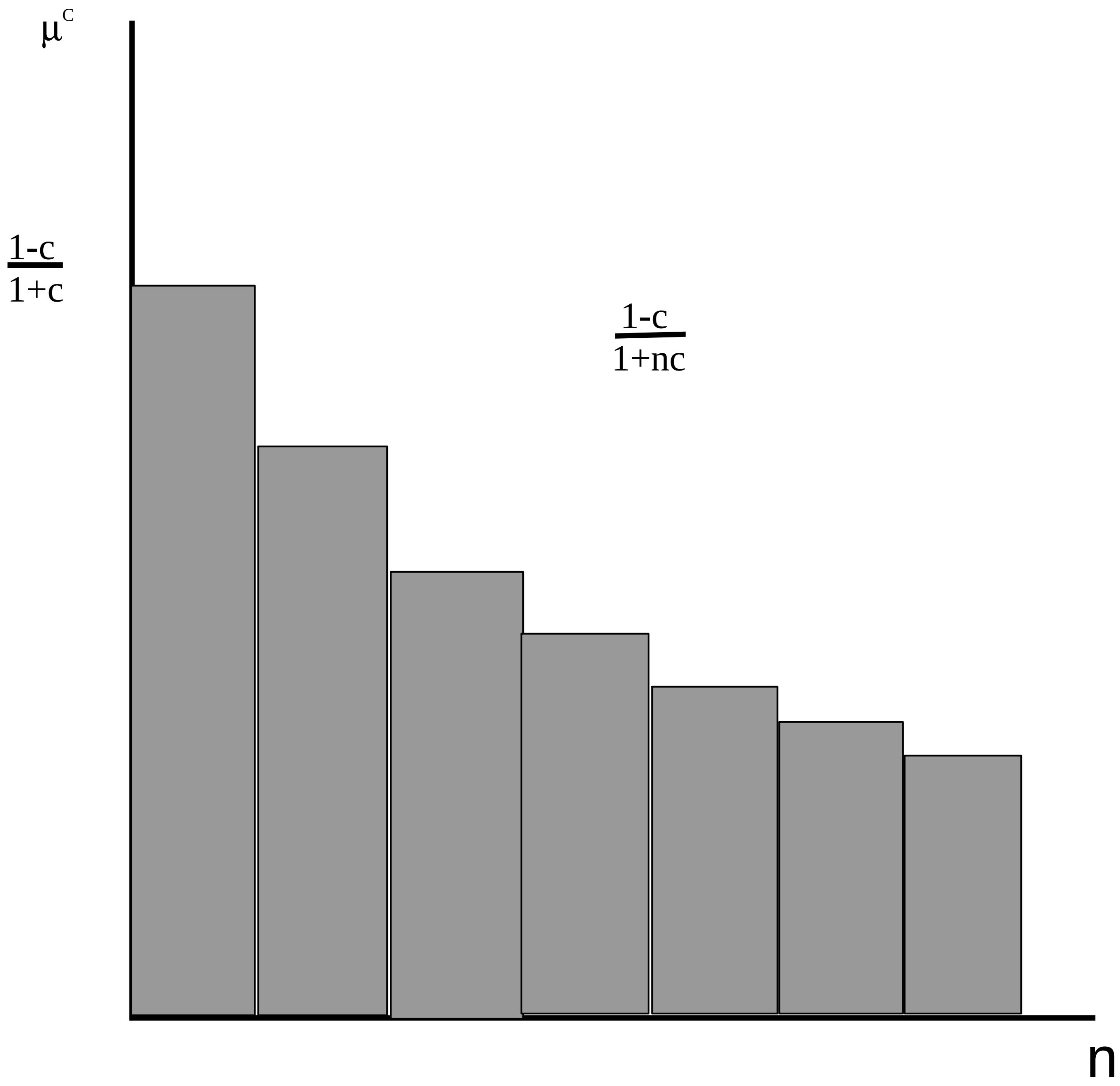


Figure 6.1: The price-cost margin and the number of firms
in the Cournot-Nash equilibrium

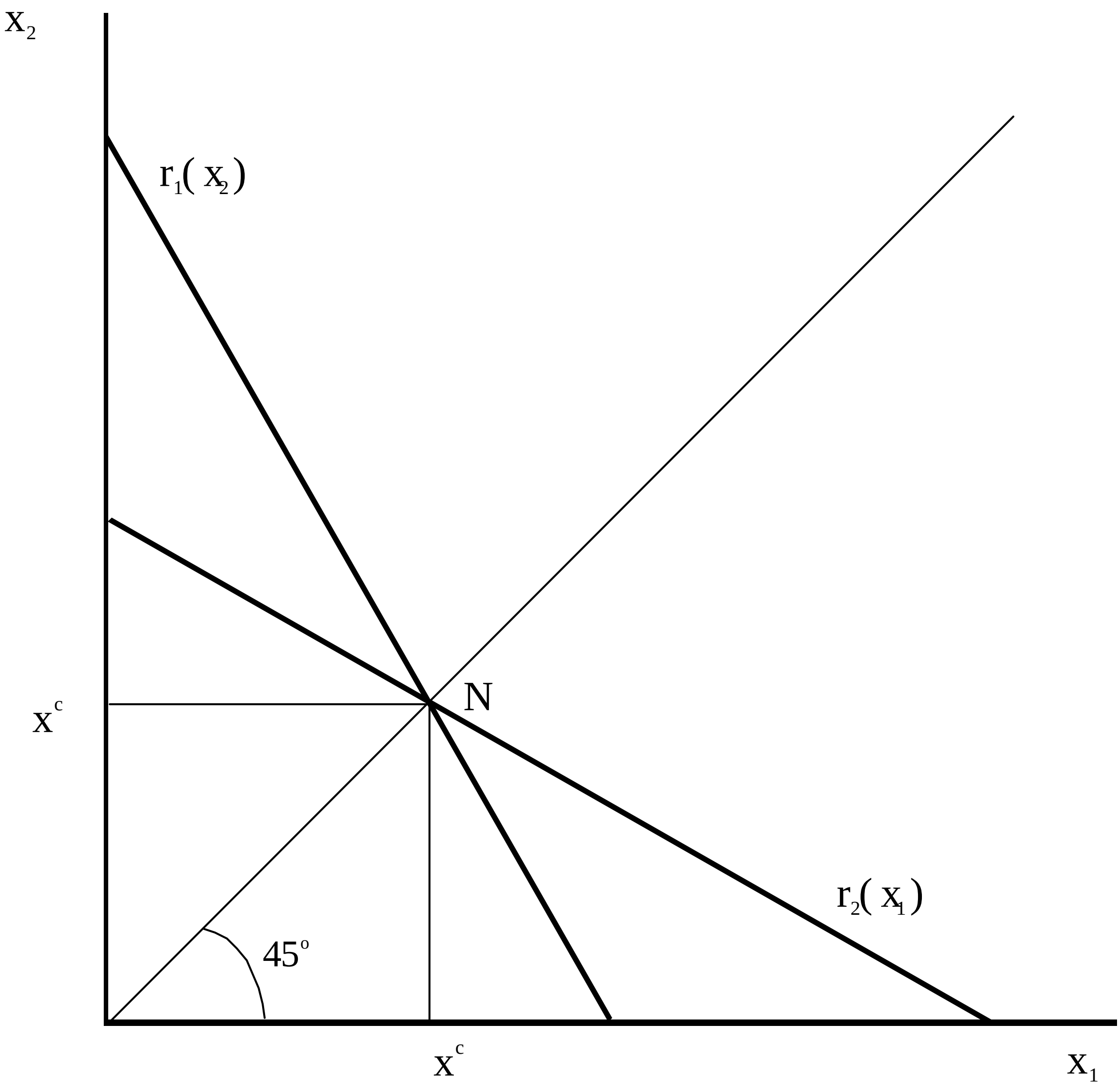


Figure 6.2. Cournot reaction functions.

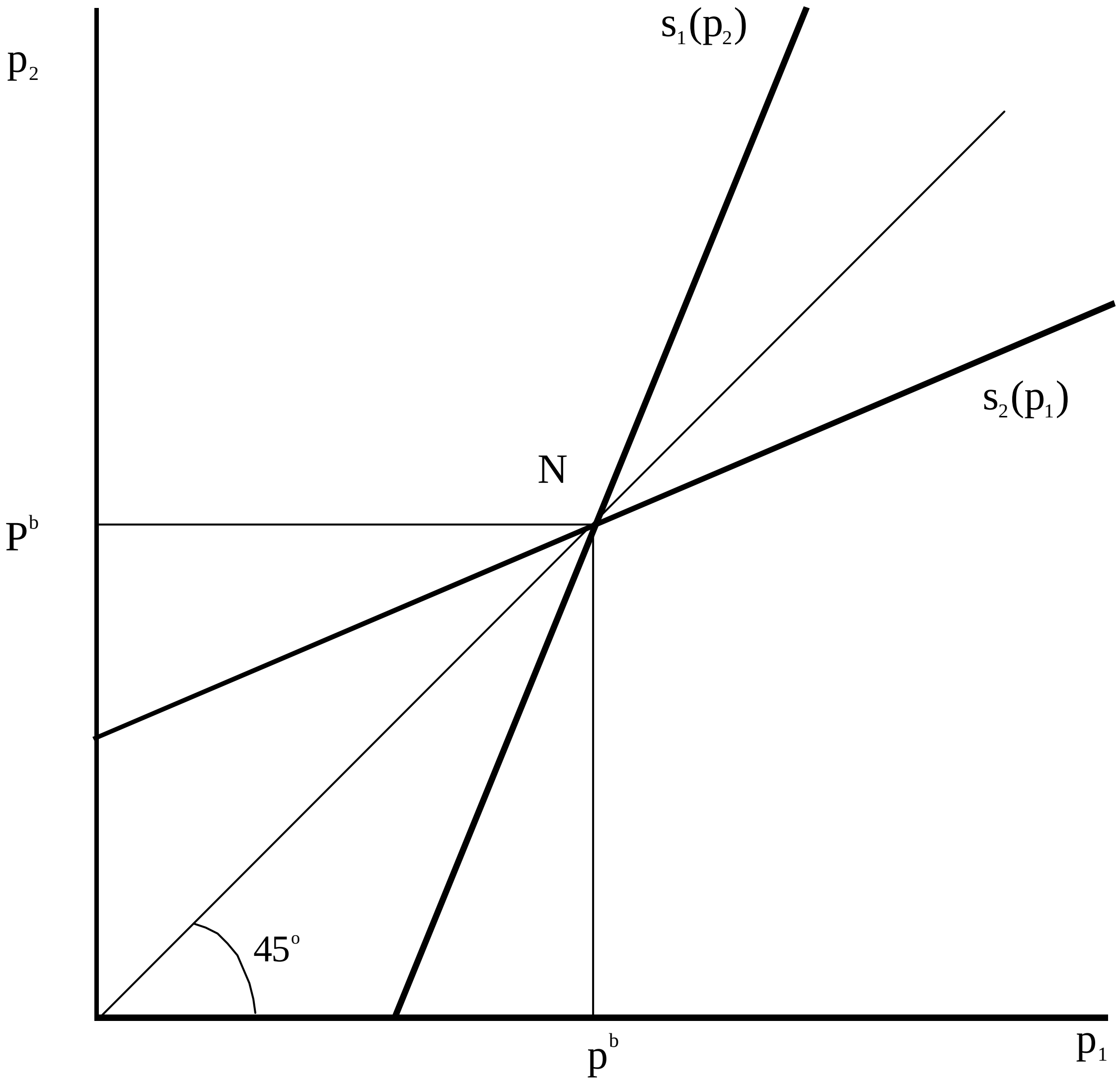


Figure 6.3.Bertrand-Nash equilibrium

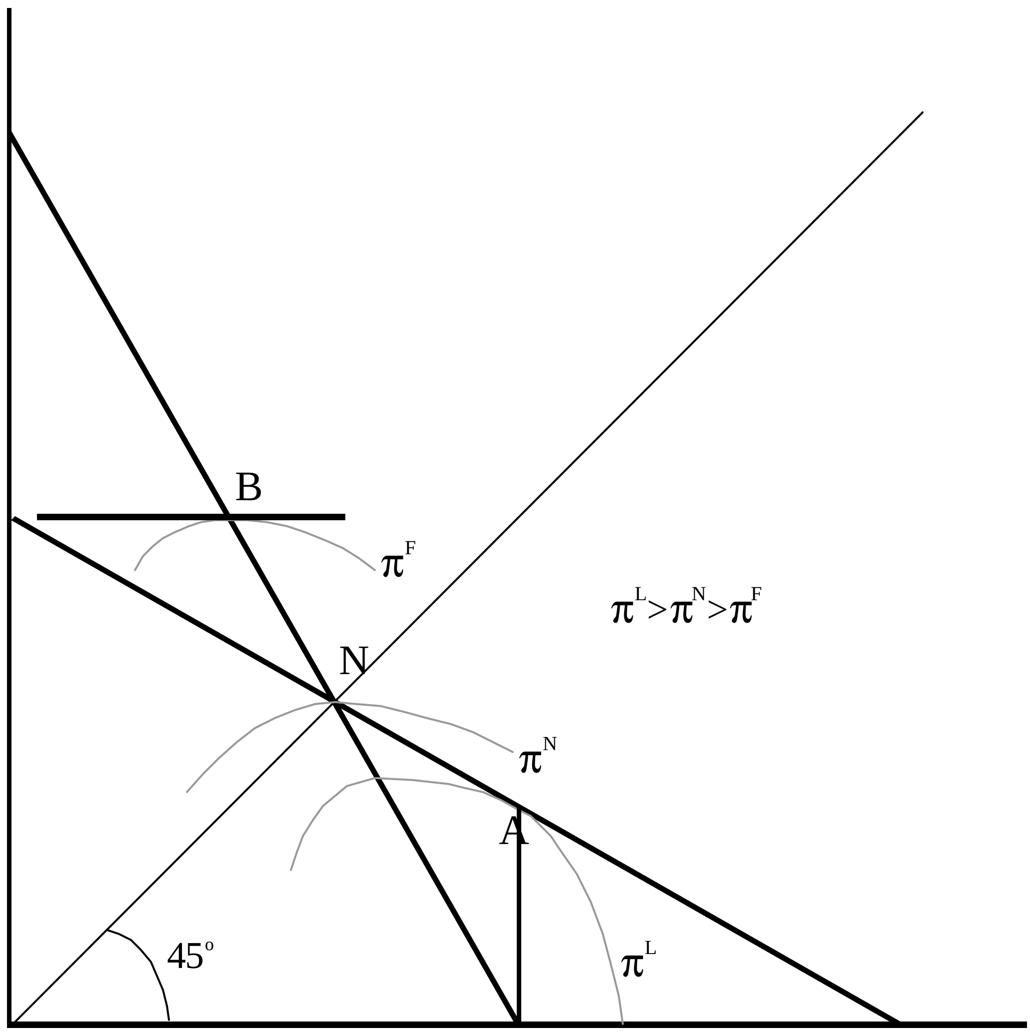


Figure 6.4. First-mover advantage in Cournot duopoly.

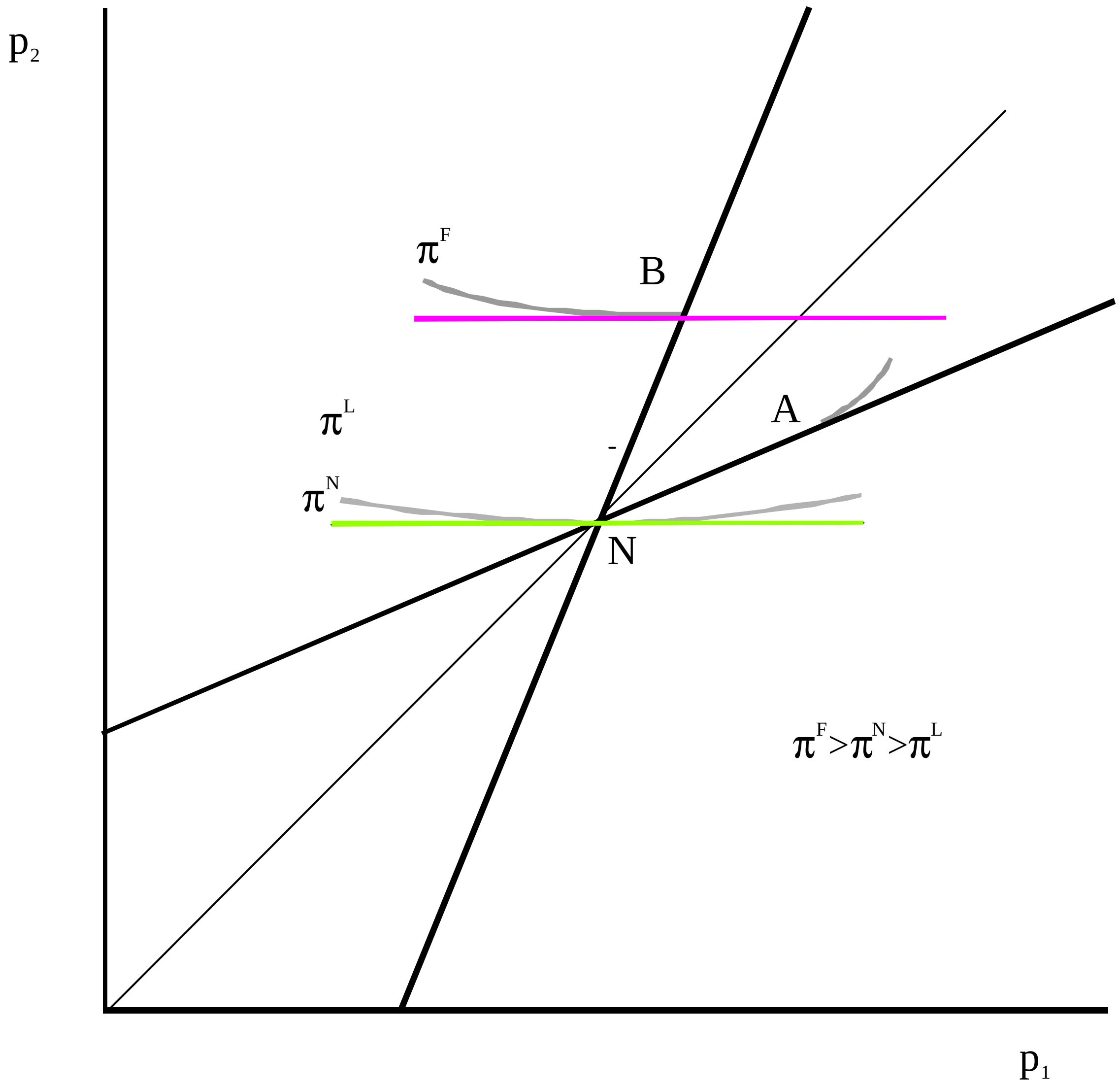


Figure 6.5 Second-mover advantage in Bertrand duopoly

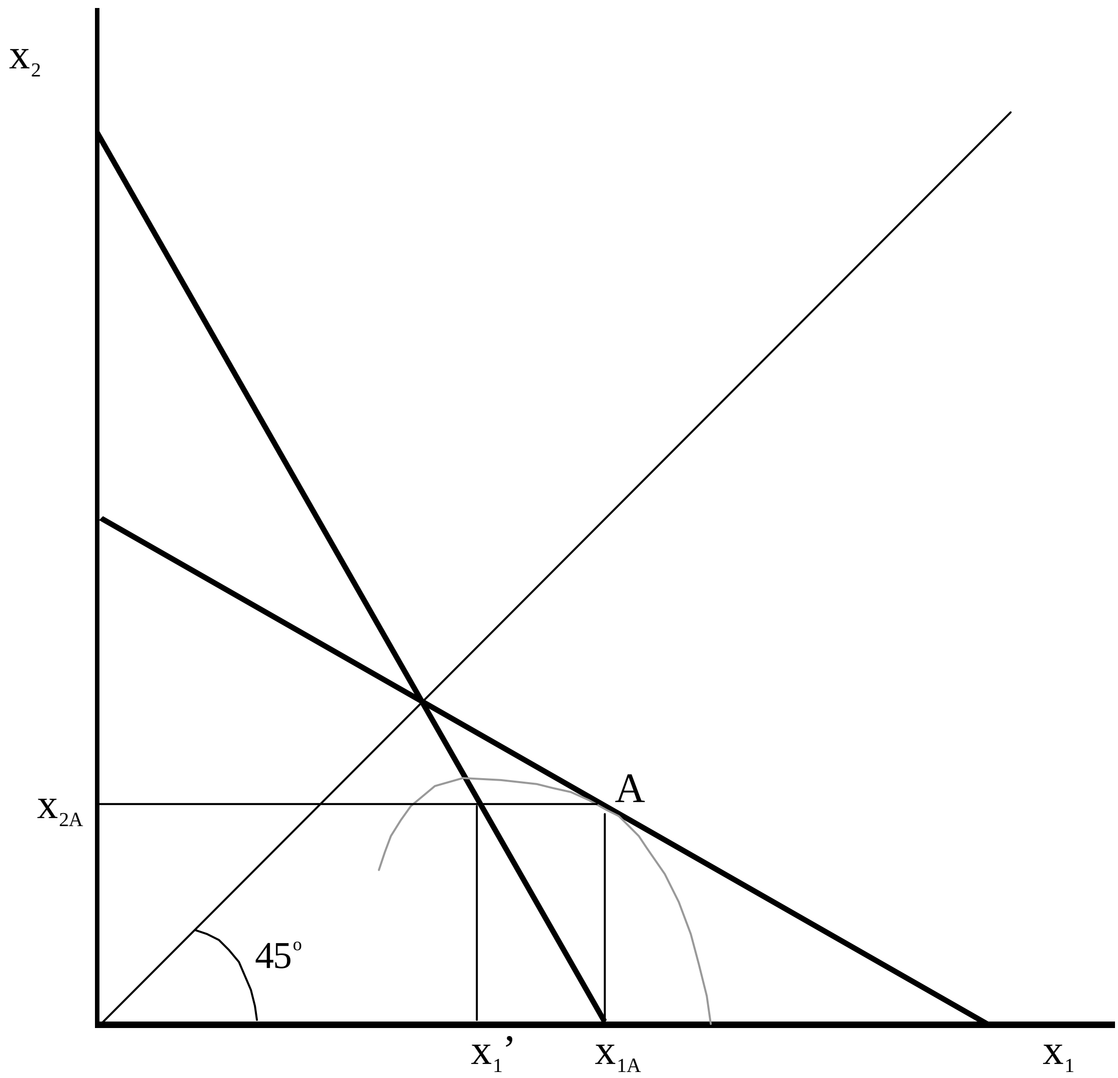


Figure 6.6. Non-credability of Stackelberg point for firm 1.

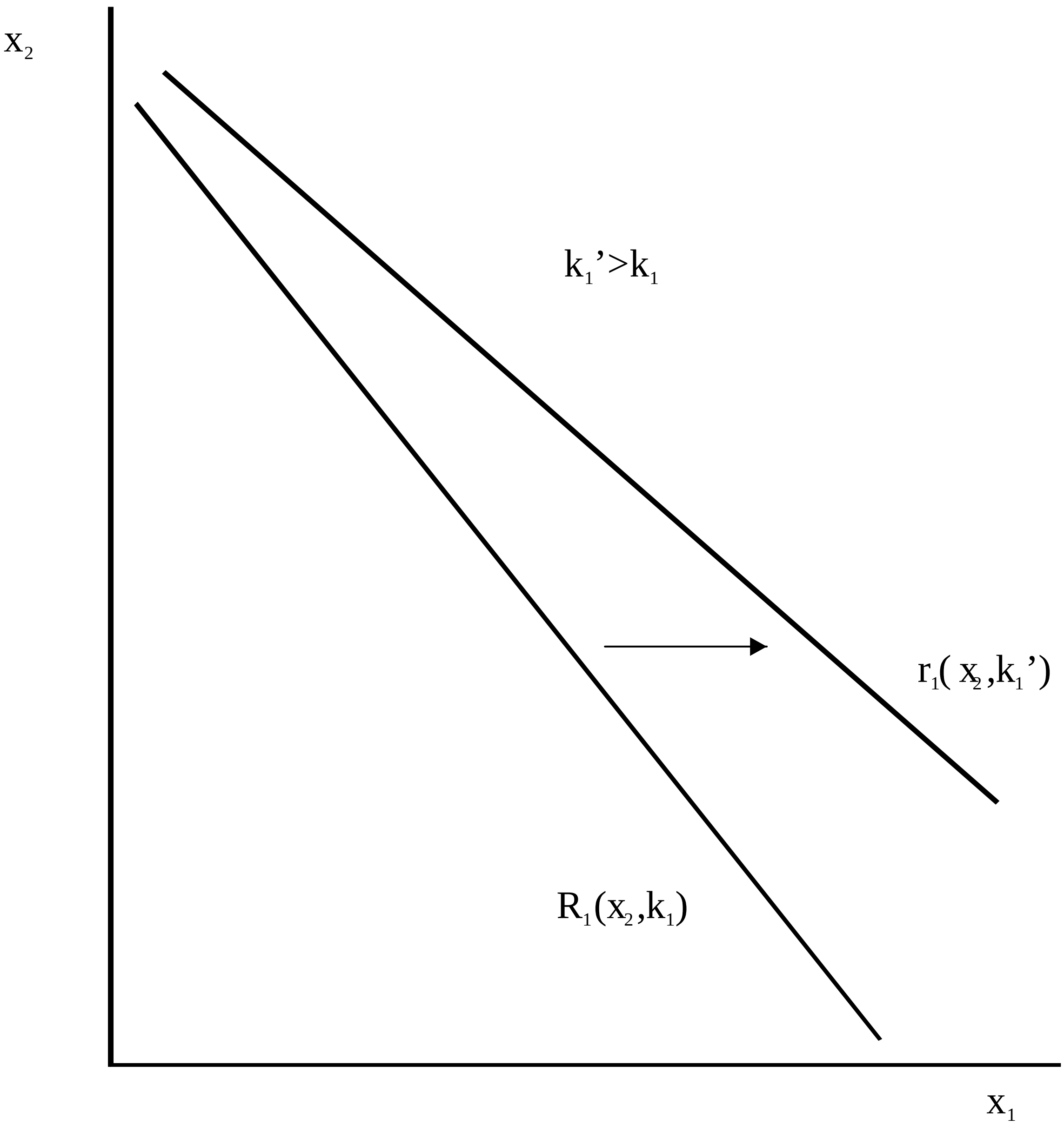
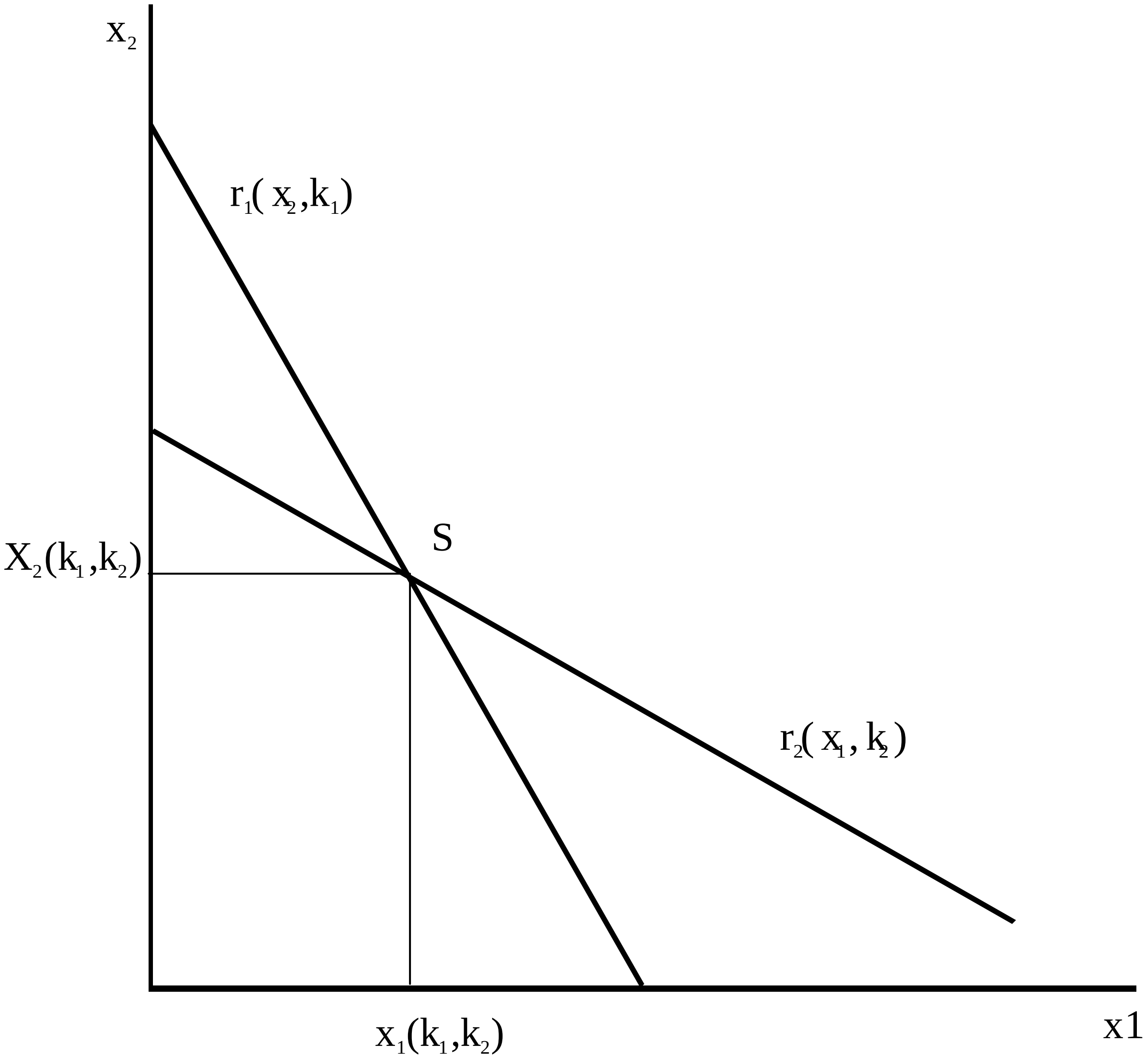
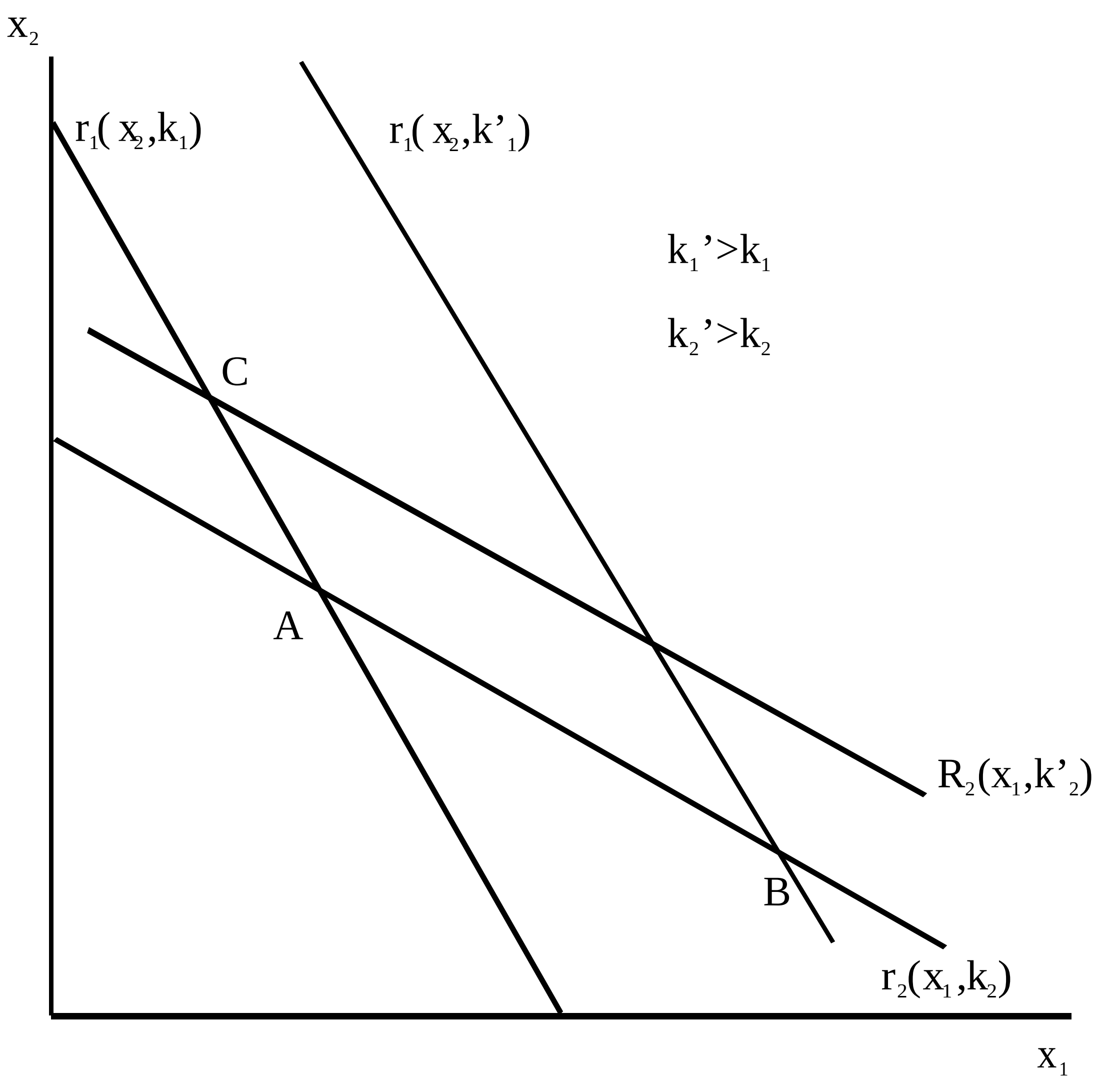


Figure 6.7. Investment shifts firm 1's reaction function out.



Figuer 6.6. Market stage equilibrium given investment k_1 and k_2.



Fugure 6.9. Market stage equilibrium and changes in investment.

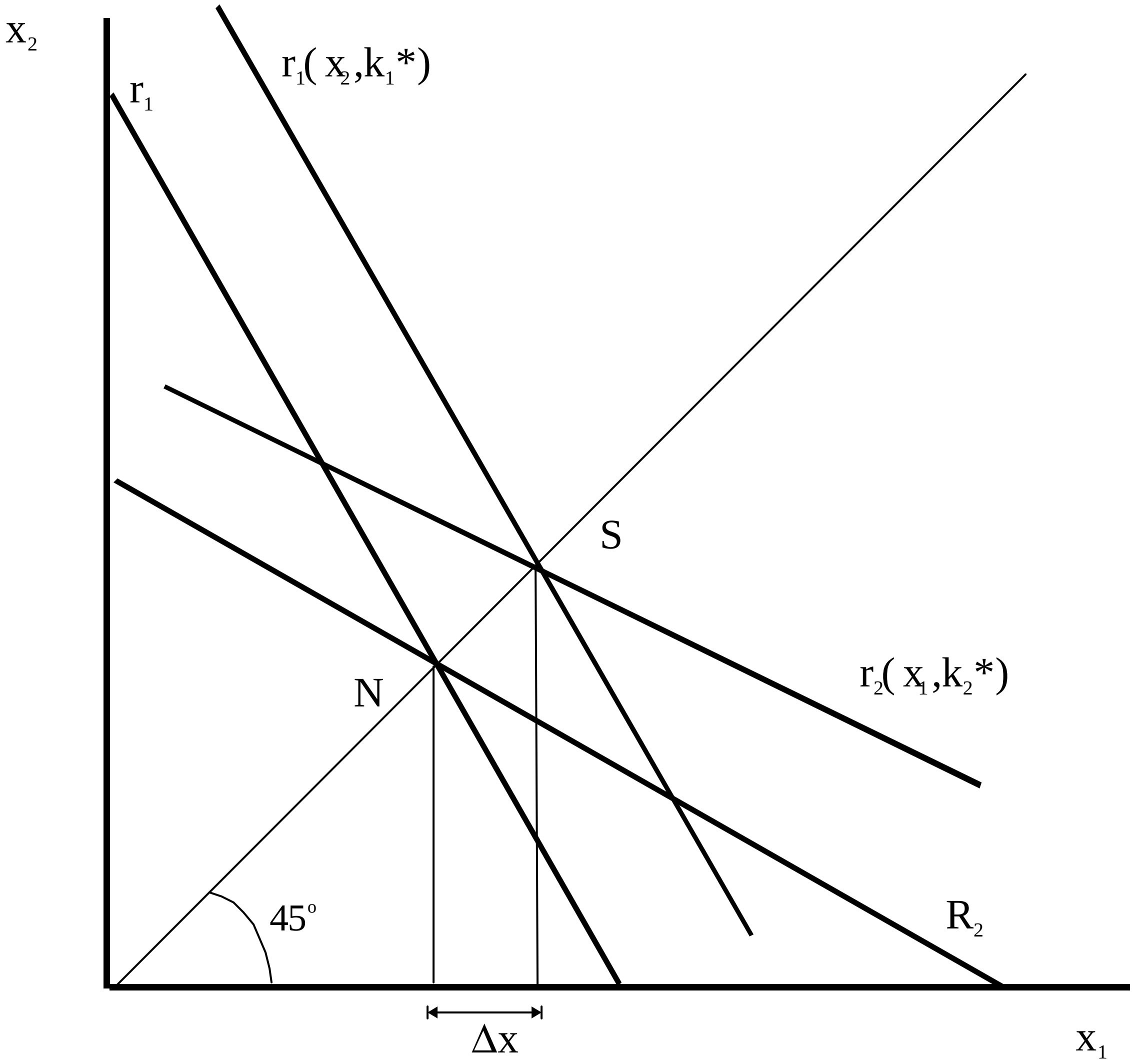


Figure 6.10. Strategic investment equilibrium: the Cournot case.
Overinvestment increases output and reduces the price.

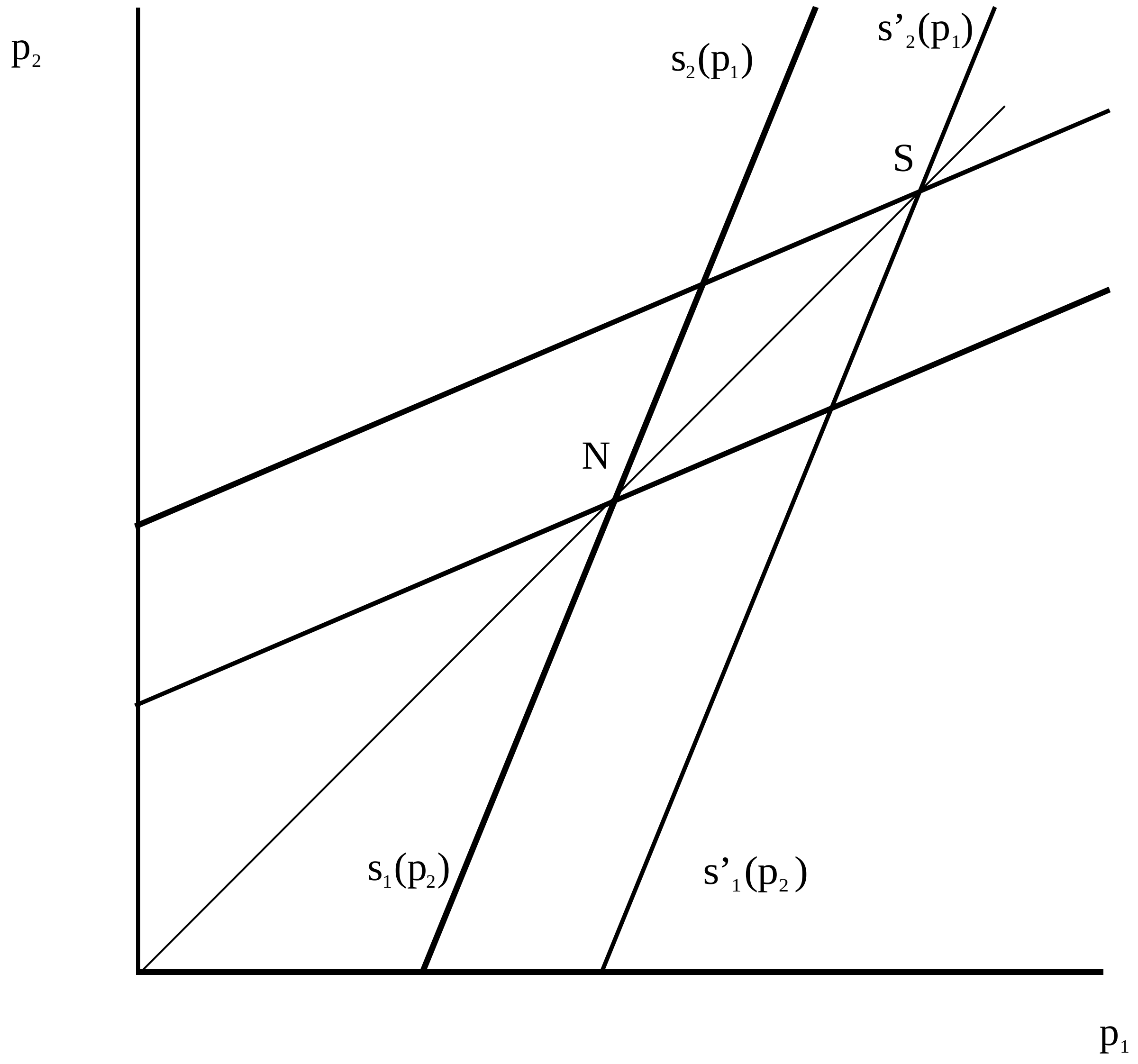


Figure 6.11. Unerinvestment raises prices - the Bertrand case.

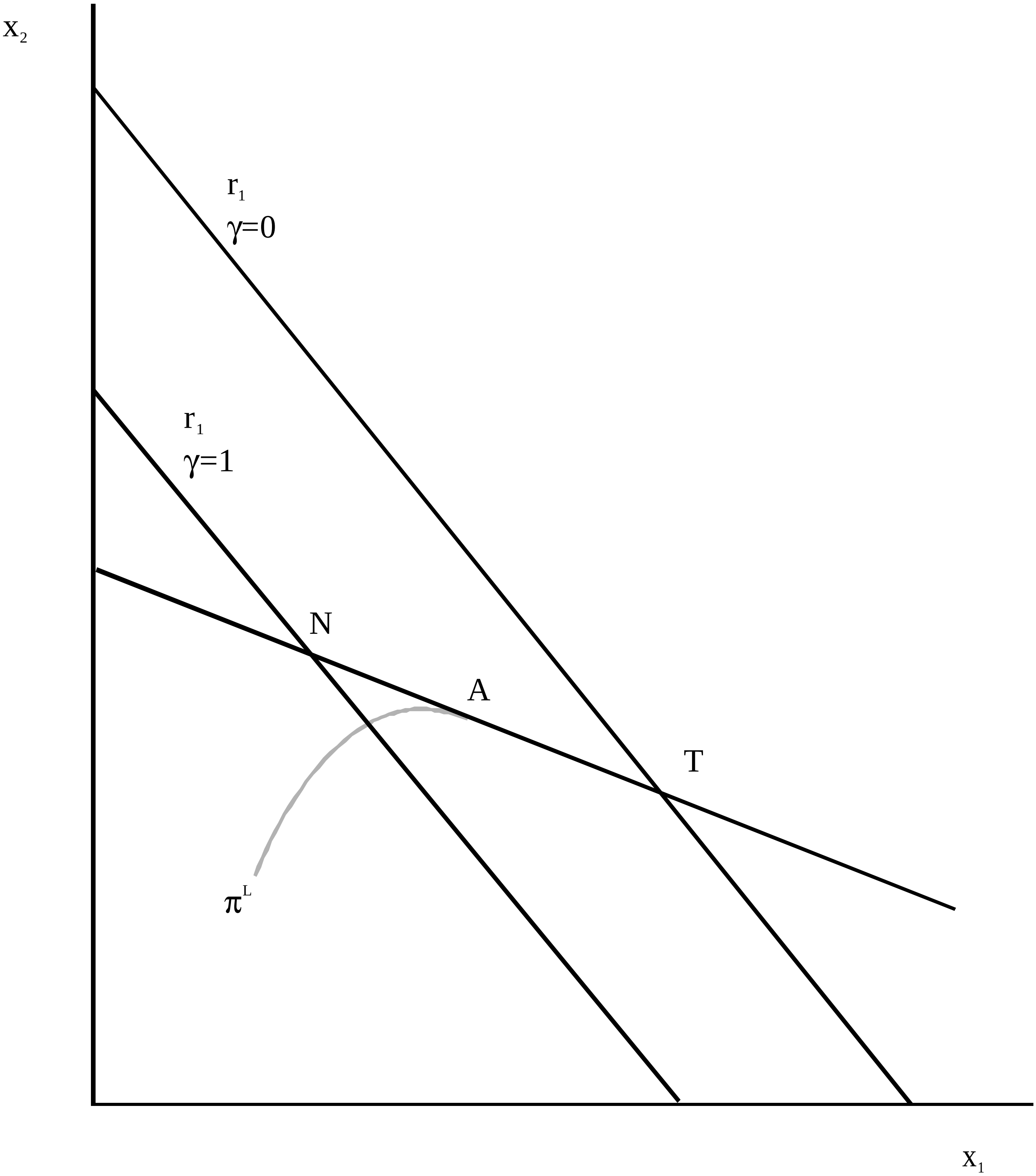


Figure 6.12. Equilibrium outcome and managerial preferences.

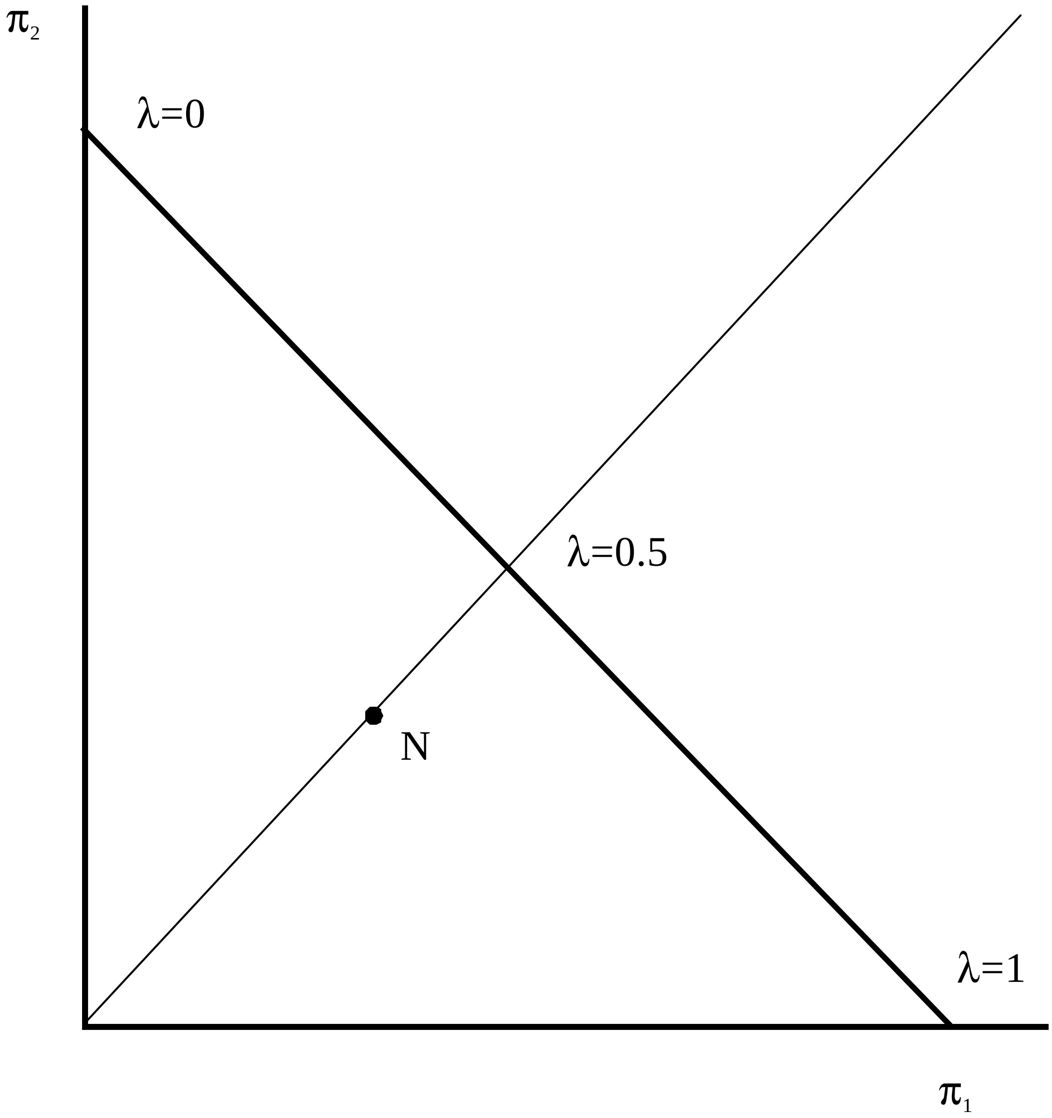


Figure 6.13. The profit frontier.