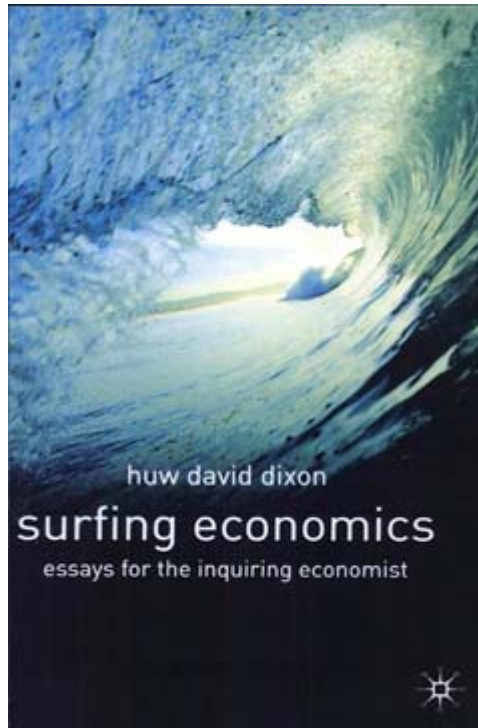


© Huw Dixon



EQUILIBRIUM AND EXPLANATION

Chapter 2 of Surfing Economics.

Originally published as: "Equilibrium and Explanation", John Creedy (ed) *The Foundations of Economic Thought*, Blackwells, 1990, chapter 13, pp.356-394.

© Huw Dixon

Chapter 2

Equilibrium and explanation

This chapter presents my personal views on the method of equilibrium in economics. The problem posed by this topic is that there is no general *concept* of equilibrium: rather, there is a *method* of equilibrium analysis that is employed in most of economics. Thus, the subject of enquiry is as diverse as economics itself. The chapter is therefore a compromise between a discussion of the general method of equilibrium itself, and an examination of its various manifestations in particular economic models. A resolution of this tension is sought by a detailed examination of the models that seem to form the core of standard undergraduate and graduate courses. This is a useful exercise since the ‘concept’ of equilibrium is born early on in a student’s career – indeed, in the opening lectures of an introductory course. Once it has been introduced, it is rarely discussed in detail again.

Since there are so many different types of economic models with diverse preoccupations, the selection reflects personal whim and current fashion. Ten or 20 years ago¹, these factors would have indicated different topics: most notably capital theory and the associated issues raised by growth theory would have been central to any discussion of equilibrium. Consideration of these issues has been omitted mainly because they were well aired at the time.

In this chapter, attempts have also been made to steer clear of technicalities. Although technicalities sometimes enable a more precise expression, they all too often embroil us in details so that we lose the more general perspective. More importantly, the technicalities often add little or nothing to the basic equilibrium concept employed. Lastly, I have sought to avoid long-winded caveats and qualifiers, leading to some sweeping generalizations. The arguments should be seen as expressing a point of view, and it is certainly accepted that there are many other possible interpretations of particular models.

At its most general, we can say that ‘equilibrium’ is a method of solving economic models. At a superficial level, an equilibrium is simply a solution to a set of equations. However, there is more to it than that. Whilst economists rarely argue over how to solve equations, they do argue over whether a particular solution

¹ Editor. This chapter was written in 1990.

represents a 'real' equilibrium or not. What is at stake is the economist's view of economic agents and the market.

Equilibrium method has come to play a central role in explanation in economics. In this it differs from other sciences, where disequilibrium states are also the object of explanation. What then is the role of equilibrium in the process of explanation in economics: how do economists explain? The remarks in this chapter are restricted to theoretical economics. Very generally, economic explanation consists of two levels. At the first level, the *microstructure*, the economist posits the behaviour of the elements of the model. In most microeconomics, the basic elements of the model consist of agents (e.g. firms and households), whose behaviour is explained and understood as some form of constrained optimization. The agent maximizes some objective function (utility, profit) subject to some constraint (budget, technology). This constrained maximization is in effect the model of individual economic rationality. In other cases, the microstructure might simply take the form of directly postulating a behavioural relationship – as for example in Keynes's consumption function. At the second level, the *macrostructure*, the individual elements of the model are put together. (The method of equilibrium consists precisely in putting together the elements in a consistent manner in solving the model.)

This dual level of explanation is best exemplified by the most common equilibrium employed in economics: the competitive equilibrium. The microstructure consists of households who maximize utility subject to a budget constraint, and firms that maximize profits subject to a technology. The microstructure is summarized by demand and supply curves. At the macro level, we put these elements together using an equilibrium concept: in the case of competitive equilibrium, we have trade occurring at the price which equates demand with supply. Actual trades equal desired trades, and demand equals supply. The theory of the competitive market (price theory) can then be used to explain prices by relating them to individual behaviour through the market equilibrium. Thus, for example, changes in the cost of inputs can be seen as altering firms' supply decisions, and hence shifting the supply curve. This will then lead to an alteration in the equilibrium price.

Equilibrium thus plays a central role in the enterprise of explanation. Parts of the model are put together and, through the application of a particular equilibrium concept, the model comes to life. Once alive, economists see what it looks like, examining the model for properties of interest (is it Pareto optimal? How does it

compare with other models?) Sometimes the models are made to ‘dance’ through the method of comparative statics, twitching from one position to another. The economist has real or imagined properties he wishes to explain: if the model displays them, he exhibits the model as an explanation. The crux of the explanation is: ‘everyone is doing as well as they can (microstructure): when everyone is doing as well as they can, X happens (equilibrium or macrostructure).’

This is seen as explaining phenomenon X. It is rather as if an inventor has an idea, constructs a machine and proudly exhibits its performance. For the economist, however, the machine remains an idea on paper. (A notable exception is Phillip’s water model of the Keynesian income-expenditure model. This, however, is still a representation.)

Given the role of equilibrium in the process of explanation, what can be said of it in general? Perhaps the main discussion of ‘equilibrium’ occurs for most economists in introductory textbooks and lectures. This tends to be geared to the demand-supply model and/or the Keynesian income-expenditure model. We shall discuss these in detail in the next section. However, three basic properties of equilibrium in general are proposed.

- P1, The behaviour of agents is consistent.
- P2, No agent has an incentive to change his behaviour.
- P3, Equilibrium is the outcome of some dynamic process (stability).

These properties are illustrated for the demand-supply and income-expenditure models in table 2.1.

Table 2.1 Equilibrium Properties

	<i>Demand-supply</i>	<i>Income-expenditure</i>
P1	Demand equals supply	Income equals expenditure
P2	Actual trades equal desired trades	Planned expenditure equals actual expenditure
P3	Tâtonnement	Multiplier

The importance of these properties is by no means equal or uniform. In particular, the stability property P3 is often played down, since it is almost impossible to provide a

coherent account of stability in economics. P2 is the key to constructing convincing economic equilibria: if a rival economist can point out that some agent can do better than he does, then the equilibrium model presented is cast into doubt. It is thus necessary to look more closely at how equilibrium consists in agents' having no incentive to alter their behaviour. In section 2.1 the three basic equilibria employed in economics are examined. Particular issues are discussed in subsequent sections: disequilibrium analysis (section 2.2), information (section 2.3) and time (section 2.4).

2.1 The Three Crosses

In this section we examine three different equilibria which are central and paradigmatic to economic analysis. First, we shall explore the competitive equilibrium, which forms the foundation of price theory and dominates the syllabus of most undergraduate economics courses. Second, we shall consider the income-expenditure equilibrium which forms the foundation of macroeconomic analysis, giving as it does the notion of the multiplier. Thirdly, we shall consider the Cournot-Nash equilibrium: this is the standard model of oligopoly and is chosen to represent the game-theoretic approach to equilibrium. Each of these equilibria has a central role in the teaching of economics: each is represented by a simple two-dimensional diagram in which the equilibrium is represented by the intersection of two functions (invariably drawn as straight lines in textbooks). Hence each is a cross: borne by the student and named after its creator (or some approximation thereto). The competitive equilibrium has two lines, one sloping up and one down, sometimes called the Marshallian cross; the income-expenditure equilibrium has two upward-sloping lines and is sometimes called the Keynesian cross: the Cournot-Nash equilibrium has two downward-sloping lines and for consistency is called here the Cournot cross. The three crosses are depicted in figure 2.1.

Figure 2.1 Three crosses.

As we shall see, although these three equilibria lie at the centre of the discipline of economics, they are very different, and in some sense contradictory, or at least incommensurable.

2.1.1 The Marshallian Cross

“I saw the best minds of my generation destroyed by madness,
starving hysterical naked,
dragging themselves through the Negro streets at dawn looking for
an angry fix,
angelheaded hipsters burning for the ancient heavenly connection
to the starry dynamo in the machinery of night...”

Allen Ginsburg *Howl*.

The idea of competitive equilibrium stems from the vision of the market acting as an invisible hand, the price mechanism bringing into balance the two sides of the market – demand and supply. As a formal idea it surfaces in its modern form in Marshall’s *Principles* (1890). The equilibrium is represented in figure 13.2. The demand curve D slopes down; it represents the amount that households would like to buy at a given price. Thus at price P' demand is X' . The demand curve is derived under the assumption that households are *price-takers*: they treat prices as exogenous, unaffected by the action of any individual household. Under the assumption of price-taking behaviour, households have a linear budget constraint, from which they ‘choose’ the utility-maximising combination of goods. The demand curve in a particular market represents the utility-maximising quantity of the goods as the price varies (holding other prices etc. constant under the *ceteris paribus* assumption).

Figure 2.2 Supply and demand.

The supply curve S is also derived under the assumption of price-taking behaviour: it represents the quantity that firms wish to supply at each price. Profit-maximising firms will choose the output which equates marginal cost with price; hence the supply function is simply the summed marginal cost functions of firms in the market, and is upward sloping due to the assumption of diminishing marginal productivity of labour or diminishing returns to scale. The demand and supply curves represent the microstructure of the market.

In figure 2.2 the supply and demand functions are put together. The competitive equilibrium occurs at the intersection of the two lines, with resultant price and quantity (P^* , X^*). Why is this seen as an equilibrium? The argument was outlined above in terms of the three properties P1-P3. At the competitive equilibrium the

amount demanded equals the amount supplied (P1); desired trades equal actual trades (P2); at any non-equilibrium price there will be excess supply or demand which will tend to lead to a movement in price towards equilibrium – the *tâtonnement* process (P3).

Let us focus on P2, the notion that agents have no incentive to change their actions in equilibrium. What exactly do firms and households ‘do’ in a competitive equilibrium, and how might they do something else? In one sense the answer is obvious: firms and households exchange and trade; firms produce output which households consume. So it might be thought that, in evaluating the competitive equilibrium, we might consider whether firms or households might want to alter their production or consumption. For example, suppose that a firm reduces its output so that, in terms of figure 2.2, supply reduces from equilibrium X^* to X' . What will happen? This raises the general problem of specifying non-equilibrium behaviour in any model. There are at least two possibilities: first, that the price remains at P^* , and a reduction in supply simply leads to excess demand $X^* - X'$; second, that the reduction in supply leads to a rise in price to P' . In the first case, the firm’s behaviour simply leads it to produce at a point where price no longer equals marginal cost, thus reducing its profits. Hence the firm has no incentive to deviate from its competitive output. However, in the second case there will in general be an increase in profits: the reduction in output will increase price (even if only by a very small amount), which may result in an increase in profits (an certainly will result in an increase in profits if the reduction in output is small). In this case, then, the competitive outcome is not an equilibrium from the point of view of P2. This illustrates the strength of the price-taking assumption: not only is it vital to define the competitive outcome (in the sense that it defines the demand and supply functions), but it is also crucial to the notion that a competitive outcome in an *equilibrium* in the sense of P2, since it ties down out-of-equilibrium outcomes.

If we turn to P3, the competitive outcome is also seen as the outcome of the *tâtonnement* price adjustment process. There is a central paradox underlying the notion of price adjustment. How can we explain changes in price in a model in which all agents treat prices as given? One approach is to invent a third type of agent in the economy: the auctioneer. Walras based his idea of the auctioneer on the ‘market-makers’ of the French Bourse. The auctioneer is the visible, if imaginary, embodiment of the invisible hand. He has no economic involvement in the market:

no mention is made of his objectives or constraints. He just adjusts prices in response to excess supply and demand. The story is very simple; the auctioneer calls out some arbitrary prices, agents report their desired demands and supplies, and the auctioneer raises prices where there is excess demand and lowers them where there is excess supply. If and when prices attain the competitive prices and there is no excess demand, the auctioneer waves his flag (or blows his whistle) and everyone then goes ahead and trades at the competitive price: households consume and firms produce. No trade or consumption is allowed before the competitive price is reached. Hahn and Negishi (1962) suggest an alternative ‘*non-tâtonnement*’ process which allows for *trading* at disequilibrium prices, but not consumption. Otherwise, the story is similar. The competitive outcome can then be seen as the outcome of this dynamic process of price adjustment by the auctioneer. However, it is not clear what has been gained by inventing a fictional price adjustment process to justify the competitive outcome. Perhaps all this means is that, if prices respond to excess demands and supplies, then price will eventually settle down at the competitive level. But it does not tell us *why* prices respond to excess demand or supply.

Textbooks adopt a slightly different approach. Competitive equilibrium is usually introduced and explained to students in an introductory economics course. Whilst the competitive model is later developed and extended, there is often little or no thought as to what it all means.. The textbook writers therefore require an intuitive, plausible and convincing story. They argue that if price exceeds P^* then there will be excess supply (as in figure 2.2 at P'): therefore firms will want to cut prices. Below P^* , there is excess demand and prices will be bid up by consumers or raised by firms. When supply equals demand, everyone can buy or sell what they want to at P^* , and so no one will want to change price. This story may be convincing, but it is certainly not correct. Whilst it is true that prices may well change in the desired direction in response to excess demands and supplies, it is not generally true that prices will come to rest at the competitive level. Take the case where firms set prices. If the market price exceeds the competitive price there will be excess supply: firms would like to sell more, and will be rationed by some mechanism to sell less than their profit-maximising trade at that price. By undercutting the other firms by a little, any one firm can therefore attract customers from its competitors and expand sales to its desired supply at only a slightly lower price. Similarly, if there is excess demand, although the firms are able to sell as much as they want, they can increase profits by

raising prices. Thus, one might expect a situation where a non-competitive price would move towards the competitive price. There is a serious conceptual issue here: in the case of undercutting, the 'price-war' will never get anywhere – since price is usually modelled as a real variable, the undercut may be arbitrarily small (see Dixon, 1993).

However, the real issue is whether or not the competitive price itself is an 'equilibrium': will firms wish to continue setting the competitive price when and if they have arrived there? In general, the answer is no (see Shubik, 1958; Dixon, 1987). Firms will want to raise price at the competitive equilibrium (see Dixon, 1987, theorem 1). The reason is simple. At the competitive price, firms are on their supply function: price equals marginal cost. This can only be optimal for the firm if the demand curve it faces is actually horizontal. But if the firm raises its price (a little), it will not lose all its customers since, although consumers would like to buy from firms still setting the competitive price, those firms will not be willing to expand output to meet demand (their competitive output maximizes profits at the competitive price). Those customers turned away will be available to buy at a higher price. Thus if a firm raises its price above the competitive price, it will not lose all its customers but only some of them, and so it will face the downward-sloping residual demand curve depicted in figure 2.3. Since it faces a downward-sloping demand curve, marginal revenue is less than price: hence at the competitive price and output (point a), marginal cost exceeds marginal revenue and the firm can increase profits by raising price (to P' in figure 2.3). This argument rests on an upward-sloping (and smooth) marginal cost function; in the standard Bertrand case of constant marginal cost, of course, it is in the interest of firms to continue setting the competitive price. However, the Bertrand case is not at all robust, since a slight deviation from constant marginal cost destroys the competitive equilibrium.

The standard textbook story of competitive price adjustment just does not stand up to closer scrutiny. The basic problem is the contradiction between an equilibrium concept based on price taking and the notion of agents (firms or households) setting prices. Indeed, it has proven very difficult to provide a coherent account of competitive equilibrium which allows for individual agents to do anything other than choose demands or supplies at given prices. This does not mean that many minds have not put their ingenuity to solving this puzzle (see Dubey, 1982; Simon, 1984).

However, one can but marvel at the baroque intricacies needed to provide suitable clothing for the classical simplicity of the original competitive edifice.

What are we left with? We shall return again to the Marshallian cross. However, at this stage one is tempted to say that the competitive outcome does not represent an ‘equilibrium’ at all. This is in a sense surprising since, for most undergraduates, competitive equilibrium is ‘the’ equilibrium. Certainly, for a couple of generations of academic researchers, the Arrow-Debreu incarnation of the Marshallian cross held an almost ineluctable fixation.

Figure 2.3 Non-existence of competitive equilibrium

Debreu’s *Theory of Value* was published in 1959. The best minds of a few generations travelled the more arcane regions of ‘general competitive analysis’. It is the fate of each generation’s passion to seem unnatural to its successors. Perhaps the seduction lay in the panoply of technique needed to analyse the model: it was, after all, Debreu who established real analysis at the preferred language of economists. On the conceptual level, however, whether one looks at the competitive outcome as the fixed point of some mapping or the intersection of supply and demand makes little or no difference. Nonetheless, if we look closely at the Marshallian cross it seems difficult to give a coherent account of equilibrium in terms of either P2 or P3. The central contradiction is that, whilst price plays a central role in competitive analysis, no agent (excepting the fictitious) has any direct control over the price. Thus it makes little sense to say either that no one has any incentive to deviate from the competitive outcome or that the price will adjust towards the competitive price.

This does not mean that the competitive outcome is not useful, despite not being an equilibrium. It can be seen either as an ideal type, a (possibly unattainable) benchmark, or an approximation to a non-competitive market. For example, since a competitive market has the desirable efficiency property of Pareto optimality, governments may wish to make non-competitive markets behave more like competitive markets (e.g. as in the regulation of natural monopolists). Or again, under certain circumstances a non-competitive market may behave almost like a competitive market: for example, under certain assumptions the Cournot equilibrium ‘tends’ to the competitive outcome (see below) as the number of firms becomes large. This means that a Cournot market with many firms can be approximated by its

'limiting case', the competitive outcome. However, that being said, it seems wrong to view the competitive outcome as an equilibrium at all (except perhaps under certain well-specified cases).

2.1.2 The Keynesian Cross

The Keynesian cross represents the equilibrium of the income-expenditure relationship developed by Keynes (1936) and is represented in textbooks by the 45^0 line diagram developed by Samuelson (1948). The microstructure here consists of simple behavioural relationships. First, expenditures are divided into 'autonomous' (i.e. uninfluenced by income) and 'induced' (i.e. influenced by income) expenditures. Investment and government expenditure are usually seen as autonomous, consumption as induced. A basic behavioural postulate is made about the relationship between income and consumption expenditure: higher income leads to higher expenditure, and furthermore the proportion on income spend falls with consumer income. The macrostructure of the model consists of putting together these two types of expenditure (autonomous and induced) and allowing income to adjust to ensure 'consistency' between income and planned expenditure. In figure 2.4 income Y is on the horizontal axis and expenditure is on the vertical axis. Total planned expenditure is consumption plus investment $I + C(Y)$. The 45^0 line represents the locus of equality between income and expenditure. At Y^* in figure 2.4, planned expenditure equals planned income, since $I + C(Y^*)$ intersects the 45^0 line.

Why is Y^* seen as the equilibrium outcome? Again, look to introductory textbooks for the answer. Since at Y^* planned expenditure equals income, and in terms of equilibrium property P2, agents have no incentive to change behaviour. At incomes other than Y^* , there will either be an excess of planned expenditure over income ($Y > Y^*$) or the opposite ($Y < Y^*$). Given the income-expenditure identity, something has to give (this is variously explained in terms of forced saving by consumers or undesired inventory charges by firms). It needs to be noted that there is no derivation of the consumption function from some constrained maximization, and so no 'explanation' of planned consumption, which is itself a datum of explanation. Furthermore it is an *aggregate* consumption function. It is not clear how we can make sense of behaviour changes by any individual, or consumers in general, and how we might judge the effects of any change. Thus P2 is very weak here.

Figure 2.4 Income-expenditure equilibrium

However, the ‘stability’ equilibrium property P3 plays a much larger role in our understanding of the income-expenditure equilibrium. The driving force behind the equilibrium, the pump propelling the circular flow of income, is the multiplier process, which reflects the combination of a behavioural assumption (income generates expenditure through the consumption function) and an identity (income equals expenditure). One can view the equilibrium outcome Y^* as being generated by the following story. In the beginning there is just autonomous expenditure (for simplicity, say investment I). This generates income, in the sense that the act of ‘expending’ involves a transfer of money from the spender to the vendor (this is the income-expenditure identity). This income then generates consumption through the behavioural assumption that planned consumption depends upon income. As all economics students know, the end result of this infinite income-expenditure series is precisely income Y^* . This stability story is easily adjusted to allow for any ‘initial’ income.

If we consider the multiplier story underlying the Keynesian cross, we can see it is much more convincing and credible than the *tâtonnement* story. In one sense the multiplier process can never happen: it would take forever for the infinite geometric series to occur in real time and the process would run into serious problems of the indivisibility of currency. However, it is a process that we can observe, and the logic of geometric series implies that even a few iterations will move income close to Y^* . For example, when a foreign firm invests in a depressed region, it hires workers who spend money, shops open to serve them and so on. The *tâtonnement* is not observed. Whilst we see prices changing, there is no direct reason to believe that they come to rest at competitive levels. It is very important to note the contrast in emphasis on P2 and P3. Because the income-expenditure model has little microstructure to flesh out the issue of incentives to alter production or consumption, the explanatory emphasis shifts to the stability issue as embodied in the dynamic multiplier process underlying the equilibrium.

As a final comment, I do not believe that the Keynesian cross model is something of only archaeological interest in the history of macroeconomic thought. It underlies most macroeconomic models in the determination of nominal income. The behaviour

of real income is of course a different matter, which depends (usually) upon what happens in the labour market.

2.1.3 The Cournot Cross

The fact that Cournot published his *Recherches sur la Théorie Mathématique de la Richesse* in 1838 is remarkable. It pre-dates neoclassical economics by at least some 40 years, providing an analysis of duopoly that forms the basic model used in industrial organization, and is introduced as the oligopoly model in micro texts. More remarkable still, it introduced the basic equilibrium concept of modern game theory: the Nash equilibrium. For these reasons it is perhaps the archetypal model underlying current-day economics, just as Walras reincarnated as Debreu underlay the economic theory of a previous generation.

The basic idea of the Cournot-Nash equilibrium is very simple. Firms choose outputs, and the market clears given those outputs. The key step here is to invert the market demand function: rather than treating household demand as a function of price, price is seen as a function of firms' outputs. This mathematical inversion has significant economic implications. In the standard homogeneous good case, it imposes a single market price on the good. Thus there is not a separate price for each firm's output, but a common market price which each firm can influence by altering its output (see *Oligopoly theory made simple* for a more detailed analysis).

If we stick to the homogeneous goods case, let there be n firms $I = 1, \dots, n$, with individual outputs X_i , summing to total output X . Market price P is then a function of X :

$$P = P(X) \tag{2.1}$$

We can write each firm's profits U_i as a function of the outputs \mathbf{X} chosen by each firm (where \mathbf{X} is the n -vector of each firm's outputs).

$$u_i(\mathbf{X}) = X_i P(X) - c(X_i) \tag{2.2}$$

where $c(X_i)$ is the firm's cost function. A Cournot-Nash equilibrium is defined as a vector of outputs \mathbf{X}^* , where each firm's output X_i^* yields higher profits than any other

output X_i given the outputs of other firms X_{-i}^* (where X_{-i}^* is the $(n - 1)$ – vector of outputs of firms other than i). Formally, at equilibrium X^* ,

$$U_i(X_i^*, X_{-i}^*) \geq U_i(X_i, X_{-i}^*) \quad (2.3)$$

for all feasible outputs X_{-i} (usually any positive $X_i \geq 0$). There may of course be multiple equilibria or no equilibria: however, we shall proceed as if there is a single Cournot-Nash equilibrium.

The Cournot-Nash equilibrium is therefore almost completely defined in terms of equilibrium property P2. At equilibrium, no firm has an incentive to change its behaviour *given* the behaviour of others. Unlike in the competitive or Keynesian cross equilibria, what happens if one agent deviates from equilibrium is precisely defined. In the Cournot case, there is a function relating the outputs chosen by each firm to their profits (equation (2.2) above). In general game-theoretic terminology firms choose strategies (output) to maximize their pay-offs (profits), given the strategies of other firms. The Nash equilibrium is central to non-co-operative game theory, and its use is spreading through economics as it evolves beyond more traditional competitive or macroeconomic frameworks. The attraction of the Cournot equilibrium is that it is self-enforcing, since no one has an incentive to defect from it. Furthermore, if everyone expects a Nash equilibrium to occur, they will play their Nash strategy.

This is illustrated by the Cournot cross. The Cournot-Nash equilibrium is usually taught using the concept of a ‘reaction function’ (or, as others prefer, a ‘best-response’ function). Each firm’s reaction function gives its profit-maximizing output given the outputs of the other firms. In the case of duopoly, firm 1’s reaction function $X_1 = r_1(X_2)$ is derived by solving

$$\max_{X_1} P[(X_1 + X_2) - c(X_1)] \quad (2.4)$$

This tells us firm 1’s best response to any output that firm 2 might choose. Similarly, for firm 2, $X_2 = r_2(X_1)$. Under standard assumptions, each reaction function is downward sloping with an (absolute) slope less than unity. In plain English, if firm 1 were to increase its output by one unit, the other firm’s best response would be to

reduce its output, but by less than the initial increase in firm 1's output. The Cournot cross is depicted in figure 2.5. The Cournot-Nash equilibrium occurs at point N, where the two reaction functions cross. Only at N are both firms choosing their profit-maximizing output given the output of the other firm.

Figure 2.5

What of the issue of stability in the Cournot model? The usual textbook story is that, starting from some disequilibrium position, the firms take turns to choose outputs. At each step, the firm chooses its output to maximize its profits given the output of the other firm (i.e. it moves onto its reaction function). In terms of figure 2.6, starting from point a we follow the arrows: firm 1 moves first to its reaction function, firm 2 moves to its reaction function and so on. This process will 'converge' to the equilibrium at N: although the firms will never reach N, they will get closer and closer (in mathematical terminology, the outputs converge uniformly to N but do not converge pointwise). An alternative adjustment process that is harder to depict is simply to have the firms simultaneously adjust to each other's output for the previous period. Assuming that technical 'stability' restrictions are met by the reaction functions, the Cournot-Nash equilibrium can be seen as the outcome of some dynamic process (P3).

Figure 2.6

However, rather like the *tâtonnement*, the Cournot adjustment process lacks credibility. The crucial weakness is that, at each step, the firms behave myopically: they choose their output to maximize their current profits given the output of the other firm, but ignore the fact that the process specifies that the other firm will adjust its output as given at the Nash equilibrium N. Suppose that firm 1 alters its output to X' in figure 2.7. What would firm 2's best response be? One is tempted to say that firm 2 would move along its 'reaction function' to X'_2 . However, this will not be so in general if firm 2 envisages firm 1 making a subsequent response (since X'_2 is the best response treating X' as given). Thus the issue of how firms respond to each other is rather convoluted: each firm's response depends on how the other firm will respond to it, which depends on what the other firm thinks that the first firm thought ...

Fig 2.7

However, if we alter the firm's conjectures, we not only alter the process of adjustment, but also the equilibrium itself. The reason for this is that in taking output

decisions firms will take into account the other firm's response, rather than treating the other firm's output as given as in the Cournot-Nash equilibrium.

The nature of conjectures can be very general, allowing for the initial position and the size of the change. For example, firm 1's conjecture about firm 2's output could give X_2 as a function K of X_1 and the initial situation (X_{10}, X_{20}) :

$$X_2 = K(X_1, X_{10}, X_{20})$$

A much simpler (and more common) form of conjecture is to restrict firms to a specific form of conjecture – namely a proportional response which is invariant to initial position. This restricts firms to constant conjectures z about dX_j/dX_i , called 'conjectural variations'. In a symmetric equilibrium with two firms and a homogenous product this results in the price-cost equation

$$\frac{P - c'}{P} = -\frac{1 + z}{ne}$$

where z is the conjectural variation, c' is marginal cost and e is industry elasticity of demand. If $z = -1$, each firm believes that industry output is constant, since the other firm reduces its output to offset an increase in firm 1's output exactly. In this case $P = c'$ – price equals marginal cost – and we have the competitive outcome. If $z = 0$ we have the Cournot outcome (firms believe the other's output is constant). If $z = +1$ then the price-cost margin equals $-1/e$, which is the collusive or joint profit-maximizing outcome. Industrial economists often use the conjectural variation model of Cournot oligopoly because just a single parameter (conjectural variation z) can capture the whole range of competitive behaviour from perfect competition ($z = -1$) to collusion ($z = +1$). The crucial point to note is that the nature of the equilibrium depends crucially on the conjectures firms have about out-of-equilibrium behaviour.

Thus the issue of whether or not firms have an incentive to deviate from equilibrium or not (P2) depends on how they conceive of disequilibrium. This causes problems for stability analysis in the Cournot model: if we allow firms to be aware of the fact that they will respond to each other out of equilibrium, then in general there is no reason for them to treat each other's output as given in equilibrium. The 'myopia'

of the adjustment process is crucial for its convergence to the Cournot-Nash equilibrium at N rather than at some other point.

Current sentiment, taking its cue from game theory does not view the conjectural variations approach with esteem. There are perhaps two main reasons for this. First, the Nash equilibrium is seen as the only sensible equilibrium concept to employ. There is no question of being 'out of equilibrium': rational players would never play any but the Nash equilibrium strategies. Any other choice of strategies would involve one firm or the other being off its best-response function and hence able to do better. No players who recognised each other as rational would play anything but the Nash strategy. Second, on the issue of adjustment, it is currently popular to argue that the conjectural approach attempts to capture dynamics in a static 'one-off' model. If we want to understand the responses of firms to one another over time, we need to have a fully specified dynamic model in which time is explicitly present. In game-theoretic terms, this means dealing with (finitely or infinitely) repeated games, which allow firms to react to (and anticipate) each other. In the conjectural approach, the firm simply considers a range of simultaneous possibilities (different values of its own output with the corresponding conjectures about the other's output).

Under the influence of modern game theory, many economists would reject the importance of traditional 'stability' analysis, and with it equilibrium property P3. Equilibrium is to be purely defined in the form of consistency (P1) and in terms of the incentive to play equilibrium strategies given the other player's strategies (P2). There is a price to be paid for this approach of course, i.e. the notion that we must at all times be in equilibrium. This creates some problems, particularly in repeated games which we shall discuss in section 2.4.

Comparing the Marshallian and Cournot crosses, we can see that, whereas the Marshallian cross is ill conceived and defined as an equilibrium, the Cournot cross is well defined in the sense that it is clear on what firms do and how the incentive to deviate from equilibrium is specified. This has led many economists to see the Cournot-Nash approach as a way of rationalizing the Walrasian approach. The basic idea is to see the Walrasian equilibrium as the limit of the Cournot-Nash equilibrium as the number of firms in the industry becomes infinite or, - more precisely, - as the market share of each firm tends to zero. In this sense, the Walrasian outcome can be seen as an approximation to the 'true' Cournot-Nash equilibrium if there are many firms. The argument here is simple. The Walrasian equilibrium is based on the

notion of price-taking behaviour, which means that firms treat marginal revenue as equal to price. Under Cournot competition marginal revenue is less than price since, as the firm increases output, the price falls. The extent to which the firm's marginal revenue is less than price depends upon the firm's elasticity of demand e_i which in turn depends upon its market share s_i and the industry elasticity e :

$$\begin{aligned} \mathbf{e} &= \frac{P}{X_i} \frac{dP}{dX} \\ &= \frac{X}{X_i} \frac{P}{X} \frac{dP}{dX} \\ &= \frac{1}{s_i} e \\ &= ne \end{aligned}$$

If the market share s_i is very small, then we would expect the firm's elasticity e_i to become very large. If the firm's elasticity becomes very large, then marginal revenue becomes closer to price (recall that price taking is often described as having a 'perfectly elastic' or horizontal demand curve). Thus the behaviour of a Cournot-Nash equilibrium with many firms will be close to that of a Walrasian equilibrium.

An Evaluation

We have explored three different equilibria which lie at the heart of economics. To what extent do they embody a common equilibrium methodology? At the most superficial level, they do. We identified the three equilibrium properties P1-P3 that are often put forward and which seem to encapsulate the general view of equilibrium. All the equilibria possess these three properties in some sense. However, when we look at the equilibria closely, we can see that there are tensions and inconsistencies between both the equilibrium properties themselves (more specifically P2 and P3) and the equilibria. The inconsistency between P2 and P3 arises because of the nature of economic explanation. The tension between different types of equilibria arises because of substantive differences in their vision of how the economic world is conceived.

Let us first address the tension between P2 and P3, a problem that arises in both the competitive and Cournot equilibria. At the heart of this tension lies the problem of

explanation. The behavioural model of agents in the microstructure gives rise to the state of the macrostructure. In equilibrium, these two are consistent. Out of equilibrium, they are not. Thus the behavioural model which defines the equilibrium may not be suitable out of equilibrium. Most importantly, it may lead agents to behave suboptimally out of equilibrium, which makes explanation of such behaviour difficult for economists. This problem occurs when we consider P2 and P3: both involve some consideration of both equilibrium and disequilibrium states. In evaluating P2, we consider what happens when agents deviate from equilibrium behaviour, so moving away from equilibrium; in stability we analyse the equilibrium state as the outcome of a sequence of disequilibrium states ('outcome' in the sense of limit, end-point or destination).

The tension between P2 and P3 arises because the equilibrium behaviour used to define equilibrium under P2 seems inappropriate for analysing stability P3. The argument can be put briefly as follows.

- 1 Equilibrium is defined (partly) by behavioural postulates X.
- 2 In analysing stability (P3), the reason that the economic system moves towards equilibrium is that agents adjust their behaviour. Thus the driving force of the move towards equilibrium is the response of agents to their mistakes.
- 3 If agents adapted their behaviour to the disequilibrium situation, the motion of the economic system would differ and, most importantly, it might lead to a different equilibrium. In particular, behaviour Y which is appropriate in disequilibrium may not be consistent with the equilibrium defined by behaviour X.

This tension is present in both the Marshallian and Cournot cross equilibria. In the competitive cases, the equilibrium behaviour employed under P2 is price-taking behaviour. This makes reasonable sense in equilibrium (agents are indeed able to buy and sell as much as they wish), but, in analysing stability, P3, the assumption of price-taking behaviour becomes silly – prices change in response to excess demand, and agents are unable to realize their desired trades. The presence of excess demands drives the *tâtonnement* process, and the presence of excess demands reflects the fact that agents are making mistakes in some sense. Were they to adapt their behaviour to disequilibrium, the resultant end-point might be different from the competitive

equilibrium. In the Cournot case we illustrated this argument using conjectural variations.

To turn to the second issue, each of the equilibria embodies a different vision of the economic world. In the competitive world we have individual agents responding to price signals which they receive from an impersonal market. They do not see their own actions influencing anything or anyone but themselves. In the Cournot game-theoretic world, although agents act independently, they see that their own pay-offs (profits, utility) depend upon the other agents' actions as well as their own (and that their own actions influence others). There is no 'invisible hand' or price mechanism to coordinate activity across the economy. Rather, each agent acts in his own self-interest, ignoring the effect of his actions on others. In the Keynesian cross model, there are no individuals as such. Attention is focused on impersonal aggregates which are driven by their own laws rather than through any very specific modelling of individual action.

2.2 Disequilibrium and Equilibrium

The analysis of disequilibrium poses great problems for economic explanation. Whilst it is not plausible to maintain that every market and agent is at all times in equilibrium, economists have very little to say about what happens out of equilibrium. This is because equilibrium itself lies at the heart of economic explanation. There is a sense in which economists cannot explain out-of-equilibrium phenomena. If we recall the structure of explanation in economics, it rests on two levels: microstructure, where agents' behaviour is specified (usually as the outcome of constrained optimizations); and macrostructure, where the parts are put together into a whole by the use of an equilibrium concept. Putting these two levels together, we can explain: 'everyone is doing as well as they can (microstructure); if everyone is doing as well as they can, then this happens (macrostructure)'. The equilibrium concept thus relates the properties of the whole (be it macroeconomy, general equilibrium system, market or firm) to behaviour and motivation of the individual agents or parts.

The problem of disequilibrium analysis is that it is not an equilibrium, so that by definition the desired behaviour of agents is not consistent and in some sense their actions are not the best they could take. In essence, explanation of disequilibrium involves the explanation of a mistake. In a theory based on formal rationality (as

opposed to procedural rationality), it is difficult to explain mistakes. With a procedural model of rationality, the explanation of a mistake can consist in showing that (in a particular instance) the procedure generating choices and actions is wrong. With a formal notion of rationality, there is no reference to the procedure: the decision is linked directly to the basic conditions of the problem, and there is no room for explaining the mistake. This difference is perhaps best illustrated by the contrast between adaptive and rational expectations in the context of the Lucas supply curve. With adaptive expectations, there is a procedure generating the expectations of agents in the economy; it may be a good or bad procedure depending on the behaviour of inflation. Whether it is a good or bad procedure, it can explain why agents may make systematic mistakes under certain circumstances (e.g. if inflation accelerates). With rational expectations, however, there can in principle be no explanation of how people make systematic mistakes. Indeed, that people make systematic mistakes would in itself be a refutation of the rational expectations hypothesis. Thus the formal notion of rationality employed by most economists is rather ill equipped to deal with disequilibrium.

This point is perhaps best illustrated by looking at disequilibrium in a competitive market. Although more of an allegory than a serious explanation of what happens, we can consider the *tâtonnement* process. Suppose that the price is above the competitive price. Supply exceeds demand, so that desired trades are 'inconsistent' (equilibrium property P1 is not satisfied). the price is falling, so 'price taking' is not appropriate. As to the issue of whether or not agents are doing as well as they can, that depends on how what goes out of equilibrium is specified. The traditional *tâtonnement* story gets around this by saying that there is no trade: in effect nothing happens. In the Hahn-Negishi (1962) story, trade occurs out of equilibrium: the Hahn-Negishi condition states that all agents have the excess demand of same sign. This means that only one side of the market is unable to meet its desired (or 'target') trades at the disequilibrium price. In terms of figure 2.2, this means that at P' demanding the good are able to obtain their desired trades, whilst suppliers will in some sense be 'rationed' (more about this later). Thus actual trades will be given by X_0 . The rationale for the Hahn-Negishi condition is that markets are 'efficient': if there is someone who wants to buy and someone who wants to sell, they will find each other. If the Hahn-Negishi condition were violated, then there could be agents on both sides of the market unable to trade. Clearly, there must be a sense in which those unable to

realize their desired trade could do better. In terms of figure 2.2, the rationed suppliers could cut their price: again, the price-taking assumption is not appropriate out of competitive equilibrium. The basic assumptions about the behaviour of economic agents (demands and supplies are derived by treating prices as given, and there are no constraints on trades) are not consistent unless they are assumptions of competitive equilibrium. Or, to put it another way, agents display no 'disequilibrium awareness' in Fischer's (1981) terminology; they behave as if they were in equilibrium even when they are not.

One response to this is to extend the model to allow agents to behave in a fundamentally different and appropriate manner out of equilibrium. For the case of competitive price adjustment, Fischer (1981) has attempted the *non-tâtonnement* process in this way. The resultant model is extremely complex and unwieldy. Fischer contemplated submitting an equation in his model to the *Guinness Book of Records* for the longest ever Lagrangean (1981 p.290n)! This is not the place to discuss whether or not this attempt was successful. However, even if successful, it would simply come close to defining a new equilibrium for when prices are adjusting to the competitive equilibrium rather than studying disequilibrium itself.

This illustrates the fundamental problem posed by disequilibrium: the explanation is either unsatisfactory or it leads to the definition of a new equilibrium. The best illustration of this is the literature on fix-price 'disequilibrium' models.

In the 1960s, workers on the reinterpretation of Keynes (Clower, 1965; Leijonhufvud, 1968) argued that Keynesian macroeconomics was incompatible with Walrasian equilibrium, and that phenomena such as rationing and the income-expenditure process arose from 'false trading' at non-competitive prices. Unless prices were seen as adjusting instantaneously to their Walrasian values, microeconomics would need to be revised to take account of trading 'in disequilibrium'. The response of economists in the 1970s was to pursue the study of *disequilibrium* by defining a new sort of *equilibrium*, fixed-price equilibrium, the main contributions being by Barro and Grossman (1971), Benassy (1975) and Malinvaud (1977). The study of fixed-price equilibria adopted the basic notion of price-taking from the Walrasian approach, but made prices *exogenous* (rather than trying them down to the market-clearing level). The basic task was to provide a consistent and coherent account of trade at 'disequilibrium prices'. In Walrasian microeconomics, agents believe that they can buy or sell as much as they wish at the

given price. Out of competitive equilibrium this is inappropriate, and so a new economic variable was introduced – the quantity or rationing constraint. The notion of *effective demand* was specified as the demand which (in the household's case) maximizes utility *given* the quantity constraints that it faces in other markets. Thus if a worker is unable to sell his labour (there is an excess supply of labour), this will restrict his demand for consumption goods. Similarly, if a firm cannot sell all it would like to at the current price, then its demand for labour will be influenced. This is the essence of the 'spillover' effect: if an agent is unable to realize his trades in one market, it may affect his demand or supply in other markets. In the eyes of the 'reinterpretation of Keynes' school, this was the very essence of the multiplier process and the income-expenditure feedback, which meant that quantities (rather than just prices, as in the Walrasian case) entered into demand functions.

By introducing a rationing regime into the market process, and allowing rationing constraints to influence individual agents' decisions, the fixed-price approach was able to reconcile the microstructure of the model with the macrostructure. In essence, at non-Walrasian prices, demands and supplies do not equalize; agents are unable to realize their desired trades (macrostructure). Recognizing the constraints on trade, agents revise their desired trades to take these into account (microstructure).

Thus the analysis of competitive 'disequilibrium' led to the invention of a new type of 'non-Walrasian' equilibrium. The analysis of 'disequilibrium' did not lead to a genuine disequilibrium analysis. Rather the logic of economic explanation led to the generation of another equilibrium concept. In this case, fix-price equilibria are a *generalization* of Walrasian equilibria: a Walrasian equilibrium is merely a fix-price equilibrium where agents face no (binding) rationing constraints.

Given the real disequilibrium analysis is to some extent incompatible with standard economic explanation and rationality, to what extent is economics possible without equilibrium? To see what economics looks like without equilibrium, it is salutary to look at one of the few economic models to reach textbook popularity which did not employ the equilibrium method. The example I have chosen is Kaldor's growth model (see Jones, 1975, pp. 146-9, for a concise exposition of Kaldor, 1975). The central issue in Harrod's growth model was the possible divergence of the 'warranted' growth rate (which equates planned saving and planned investment) from the 'natural' growth rate (determined by demography, technological progress etc.). Kaldor's microstructure consisted of the differential savings propensities out of wage and profit

income. Thus savings were influenced by the distribution of income between wages and profits. Rather than employ an equilibrium concept in his model, Kaldor used a 'stylized fact': namely that there had been full employment in post-war European economies. Kaldor argued that the distribution of income between wages and profits would adjust to maintain consistency between the warranted and natural growth rates. Thus a feature of the macrostructure was employed to tie down the distribution of income (to 'close the model'), rather than an equilibrium concept. This explanation is viewed as odd because it works backwards: rather than deducing macro properties from individual and market behaviour, it deduces the distribution of income across wages and profits from the macrostructure. Indeed, Modigliani and Samuelson went so far as to say 'If you can believe this, you can believe anything' (1966, p. 234). The equilibrium methodology is so ingrained in economists' minds that they will not be convinced by non-equilibrium explanations.

2.3 Information and Equilibrium

Until now we have been considering only 'full-information' models, where agents have a given information set including all the relevant information for them to take decisions. In the 1970s there was a blossoming of interest in exploring the implications of imperfect information for economic behaviour and equilibrium. Perhaps the earliest interest was in the 'search' models of unemployment in the late 1960s. In these models, agents have imperfect information about prices (wages), e.g. what prices (wages) a particular firm is offering. Search models of unemployment modelled the response of the unemployed to this problem: do they take the next job offered to them at a particular wage, or will they continue to search and incur the cost of further unemployment? The general solution to this problem was the 'reservation wage' rule: the unemployed continue searching until they are offered at least their 'reservation wage'. However, these were not *equilibrium* models, since there was no explanation of the initial distribution of wage offers. Why should firms offer different wages to the same worker or type of worker? However, the model was nonetheless very influential, not least in its influence on Friedman's formulation of the natural rate hypothesis (Friedman, 1968). We shall take Spence's (1974) signalling model as our archetypal imperfect information model. This had a tremendous impact at the time, and introduced the fundamental concept of a *signalling equilibrium* and the crucial

distinction between *separating* and *pooling* equilibria which has proved to be so important in subsequent years.

The standard approach to imperfect information is to presume that if agents do not know the true value of X they have some subjective probability density function for X , $f(X)$, in effect treating X as a random variable. Spence took the case of worker productivity, where there was asymmetric information so that, whilst workers knew their own productivity, firms did not. For example, let worker productivity X take discrete values, being either low or high (X_L , X_H respectively). Suppose that it is not possible to test directly for productivity. Whilst the employer might not know the actual productivity of an individual worker, it might have a subjective belief about the probability that the worker is high productivity (q) or low productivity ($1 - q$). Assuming that firms are risk neutral and minimize costs, they will be willing to offer workers their expected productivity $E[X]$:

$$W = E[X] = qX_H + (1 - q)X_L \quad (2.5)$$

The wage offer depends upon both the subjective probability q and the values (X_L , X_H). If we take the latter as given, what is the ‘equilibrium’ value of W and q ?

This raises an issue about beliefs in economic models. Given that beliefs are not ‘tied down’ by the truth (there is not full information), how do we explain agents’ beliefs? We shall consider this issue in this sections and the next. However, there now enters the notion of *epistemological* rationality. The agent is presumed to have (in some sense) the ‘best’ beliefs given the information available. ‘Best’ here usually means statistically optimal. An agent’s beliefs are then explained by saying that they are (in some sense) statistically optimal. ‘Best’ in this context means something completely different from the notion of ‘best’ in the theory of rational choice, where it means the choice which yields maximum utility given the constraints faced. Different notions of rationality are employed in explaining beliefs and explaining consumption or production decisions. The key point is that beliefs are not chosen to maximize utility. It may increase my utility if I believe that I am Napoleon: however, that is not a ‘rational’ belief. Let us first define economic rationality: it is an *instrumental* rationality in which choices are made merely as a means to an end (utility or profit). If we were to extend economic rationality to beliefs, people would choose beliefs so

as to maximize their utility. The belief that I am Napoleon might then be perfectly rational (indeed, not to believe I am Napoleon might be to involve some suboptimality). When beliefs clearly do matter in that they directly affect well-being (in some sense), economic rationality can be used to explain beliefs. However, in most economic models, beliefs do not enter the utility function of households directly. Debreu's households might be deist, pantheist or atheist: they might believe in general relativity or be creationists.

When explaining beliefs, however, a different notion of rationality is employed by economists. Let us take the example above of the firm hiring workers of unknown productivity. In equilibrium, is there any restriction to be placed on the firm's subjective probability q ? The usual constraint suggested is that *in equilibrium* the belief is 'confirmed'. In the case sketched above, suppose that the objective (i.e. population) proportion of high-productivity workers is q^* . Then we require that $q = q^*$: the subjective probability equals the objective probability. Assuming that the firm employs a large number of workers, the average productivity of workers the firm employs is

$$\bar{X} = q^* X_H + (1 - q^*) X_L \quad (2.6)$$

for the firm's beliefs are only 'confirmed' if

$$E[X] = \bar{X}$$

requiring that $q = q^*$. If the firm's subjective probability differs from the population proportion, then the average productivity of the workforce differs from what the firm expects.

If $q^* > q$, workers would (on average) be more productive, and vice versa. In the statistical sense, the firm would have an 'incentive' to revise its beliefs. If X deviated from $E[X]$. 'Incentive' here is used not in its standard economic sense but in a statistical sense that the optimal estimation of the population parameter would be different. Economists often tend to confuse the language and concepts of economic and statistical rationality. This is probably because both can be expressed mathematically as an optimization. However, the fact that in one case it is utility (or profit) to be maximized and in the other it is the likelihood (or some such statistical

criterion) to be maximized makes the two rationalities completely separate and incommensurable.

In most economic models, this incommensurability does not give rise to any incompatibility. However, I will give an example of when economic and epistemological rationality are incompatible. Let us return to the example of Cournot-Nash equilibrium which we explored in section 2.1. In figure 2.8, the Cournot-Nash equilibrium occurs at N: firm 1 is choosing its profit-maximizing output given the other firm's output X^N . In effect, firm 1 believes (or expects) firm 2 to play X^N , and chooses its best response. Furthermore, in equilibrium this belief is confirmed: if firm 1 produces X^N , it is firm 2's best response to produce X^N . However, suppose that firm 1 believed that firm 2 was producing X^{Oh} : its best response would be X^L .

Figure 2.8

However, suppose that firm 1 believed that firm 2 was producing X^{Oh} : its best response would be X^L . If firm 1 stuck to this belief, firm 2 would (eventually) produce X^F , its best response to X^L . This is the standard 'Stackelberg' or 'leader-follower' equilibrium with firm 1 the 'leader' producing X^L , and firm 2 the follower producing X^F . Is this an equilibrium? In one sense it is: firm 2 is choosing its best response to X^L ; firm 1 is choosing its best response to the output it believes firm 2 to be producing, namely X^{Oh} . The problem is that firm 1's beliefs are not correct: X^{Oh} does not equal X^F . If we require firm 1's beliefs to be correct, in addition to both firms' choosing their best response, there is only one possible outcome – the Nash equilibrium at N. In terms of its pay-off (economic rationality), the firm does better to have incorrect beliefs at (X^L, X^F) : it loses no profits by the fact that its beliefs about the other firm's output are incorrect (X^{Oh} differs from X^L), but clearly it gains profits by being the Stackelberg leader as a result ($u^L > u^N$). However, most economists would not maintain that this was an equilibrium: the firm would have an 'incentive' (in the sense of epistemic rationality) to revise its beliefs about the other firm's output. Thus the firm (it is argued) will change its beliefs about X_2 , so that the Stackelberg point (X^L, X^F) is not an 'equilibrium'. Thus despite the reduction in profits caused by the change of belief, epistemological rationality dictates that the belief cannot be maintained in equilibrium. There is thus a clash of rationalities, which are in this instance incompatible.

In this clash of incommensurable rationalities, economists let epistemological rationality overrule economic rationality. Even though the Stackelberg leader's delusion is profitable, he is not allowed to maintain his belief in the presence of clear evidence to the contrary. Thus we can add a fourth equilibrium property to our list:

P4, in equilibrium, agents have no incentive to alter their beliefs,

where 'incentive' is interpreted in the strictly epistemological sense, not the economic sense of P2. In the case of the simple example of firms hiring workers, P4 requires that subjective beliefs are 'confirmed', i.e. that $q^* = q$. In more complicated models, where X is a continuous variable, the beliefs might be subjective probability distributions, and P4 might require beliefs about mean, variance and possibly higher moments to be correct.

Having explored the new notion of rationality introduced to tie down beliefs in equilibrium with imperfect information, we can move on to examine perhaps the most important type of such equilibria: signalling equilibria (as introduced by Spence, 1974).

Suppose that we take the case where the firm hires workers of unknown quality, taking the model introduced earlier in this section. In equilibrium, we argued that the wage offer would equal the average productivity X of workers. This is called a pooling equilibrium, because both low- and high-productivity workers are pooled together (i.e. they are treated the same, and receive the same wages). In the full-information case, high-productivity workers would obtain their full marginal product X_H , low-productivity workers would obtain X_L . Thus the high-productivity workers do worse in the pooled equilibrium ($X < X_H$) than in the full-information case, and low-productivity workers do better. There is thus an incentive for high-productivity workers to 'signal' their ability to their prospective employer. Very simply, a high-productivity worker can signal his ability if he can do something which a low-ability worker is unwilling or unable to do. This signalling activity need have no direct causal relationship with the workers' productivity. All that matters is that the high-productivity workers have lower costs of undertaking the activity. It has long been recognized that education is used as a 'screening' device by employers to sift applicants.

Figure 2.9

Whilst education may sometimes have little direct or vocational content, it can be seen as enhancing the general ability of the educated. However, Spence (1974) abstracts from this, and focuses purely on the signalling element.

The essence of a signalling equilibrium is a circular relationship between the beliefs of agents and their behaviour. The essence is that the beliefs induce behaviour which confirms these beliefs. Thus, in the context of Spence's model, employers have a belief about the relationship between education and productivity (higher education is related to higher productivity); this causes firms to offer higher wages to more educated workers; this sets up the incentives workers face to obtain educational qualifications; workers' decisions about education determine the *actual* relationship (if any) between education and productivity. There is an equilibrium if the employer's beliefs are confirmed (P4 is satisfied). Figure 2.9 is a schematic representation.

The crucial feature of the equilibrium is that the *costs* of signalling (being educated) differ with productivity. More specifically, higher-productivity workers must have lower 'signalling' costs – in this case they must have lower costs of achieving a given level of education. The reason for this is that otherwise (given that the firm's wages are related to education) the low-productivity workers would also find it in their interest to obtain higher educational qualifications, so that the firm would offer them higher wages (in the belief that they were high-productivity workers). This is often called the 'incentive compatibility' constraint. The idea is that, in equilibrium, each different type of worker has the appropriate incentive to behave appropriately to its type. Thus, in the above case, the two types of worker differ by productivity. Suppose that the employer believes that low-productivity workers have no education, whilst those with y^* units of education (a B.A. degree) are high productivity. This wage offer might therefore be

$$W = \begin{cases} X_H & y \geq y^* \\ X_L & y < y^* \end{cases}$$

Assuming that there is no intrinsic value to education, workers will either choose no education ($y = 0$), or $y = y^*$ to get the higher wages. Workers will undertake the education if and only if the cost is less than the extra wages obtained. 'Incentive

compatibility' requires that it is in the interest of high-productivity workers to obtain their B.A. (for them, the cost of attaining y^* is less than $X_H - X_L$); low-productivity workers, on the other hand, would find it too costly to obtain a B.A. (the cost of attaining y^* is greater than $X_H - X_L$). In the signalling equilibrium (if it exists), the two types of workers are *separated*: the incentives are such that they reveal their true type through their behaviour. When this occurs, there is said to be a *separating* equilibrium. If the employer's beliefs were different, there could be a *pooling* equilibrium: if the employer believed that there was no relationship between education and productivity, he would offer the same wage regardless of education; no one would become educated, and thus the employer's beliefs would be confirmed.

In this type of equilibrium with asymmetric information and signalling, there is a very intimate relationship between economic incentives (P2) and the confirmation of beliefs (P4). However, it must be noted that in Spence's model the treatment of beliefs is very rudimentary, which results in multiple equilibria. The possibility of multiple signalling equilibria is easily illustrated using the educational signalling model. Following Spence, suppose that high-productivity workers have a cost $y/2$ of achieving education level y :

$$X_H - \frac{y}{2} \geq X_L \quad (a)$$

$$X_H - y < X_L \quad (b)$$

Condition (a) states that, given y , it pays the high-productivity worker to invest in education (cost $y/2$), to obtain extra income $X_H - X_L$; condition (b) states that this is not so for the low-productivity worker, who has to invest more effort and resources to obtain y . Both (a) and (b) will be satisfied if

$$X_H - X_L \leq y^* \leq 2(X_H - X_L) \quad (c)$$

Thus, if the employer's beliefs about the 'critical' level of education y^* are in the interval (c), then there will be a separating equilibrium with workers' activity so as to confirm the employer's belief. There is thus a continuum of equilibria. Is there any sense in which one can sensibly rank the equilibria?

Suppose that the employer and workers are economists and understand Spence's model. Then the employer will understand the incentive constraints (a) and (b). He

will deduce that anyone undertaking a level of education greater than $X_H - X_L$ must be a high-productivity worker: there is no way that a low-productivity worker would conceivably want to undertake a course of education that would cost more than the possible extra earnings. Thus, if we consider the equilibria with $y^* = X_H - X_L$, they involve an unnecessary cost in the form of education in excess of the minimum required. It has been argued (Cho, 1987) that these equilibria will not be 'stable'. A high-productivity worker can educate below y but above $X_H - X_L$ and still single himself out as being high productivity. Thus with 'sophisticated' knowledge about the way the model works, the equilibrium with the minimum amount of signalling necessary to separate types is the most plausible (in the above case, $y^* = X_H - X_L$).

2.4 Time and Equilibrium

The passing of time is a central feature of human experience. It plays a central role in much economic activity, since production takes time and consumers have to decide how to spread their consumption and labour supply over their (uncertain) lifetime. Yet, all three equilibria studied in section 2.1 were in a fundamental sense static equilibria. They were equilibria in a timeless world, or at most equilibria at a point in time which is unconnected to the past or future. How does our conception and evaluation of equilibria alter when we allow for time?

First, consider the Walrasian equilibrium. If we introduce time into the picture, there are two fundamentally different ways of conceiving of equilibrium. First is the Arrow-Debreu model, which sees the earth as a large market-place and world history as the working out of contingent contracts. Second is the notion of temporary equilibrium, which sees history as a *sequence* of transitory equilibria. We shall deal with these two ideas in turn.

Competitive general equilibrium theory explores the issue of the existence and characterization of competitive equilibrium in an economy with an arbitrary number of markets. For example, these might be seen as corresponding to n basic commodity types (bananas, nuts etc.). As such, the model is timeless. We can then ask how time can be brought into a timeless model. This can be done by a logical exercise of dating

commodities. Suppose that the world lasts T periods $t = 1, \dots, T$. We can call a banana at time t a particular commodity, to be different from a banana at time s (where s is not equal to t). Rather than having n markets, we will now have nT markets, corresponding to the n basic commodity types over T periods. This is depicted in figure 2.10 where time proceeds from left to right, each row representing a commodity type, and column a date.

Figure 2.10.

A particular square represents the market for a particular commodity i at a specific date t . Each market has a price P_{it} . Consumers will derive utility from the consumption of all nT dated commodities, which they will maximize subject to a budget constraint reflecting relative prices. Thus a household can as much sell an umbrella in the year 2000 for a corset in the year 1901 as it can exchange a banana for a nut in 1990.

In a certain sense, then, introducing dated commodities and their corresponding markets and prices into the abstract notion of competitive general equilibrium enables time to be included in the model. This is really just a purely logical exercise, however. To see why, just think what the Arrow-Debreu world would be like to live in! Competitive general equilibrium occurs at a list of prices – one for each market – at which demand equals supply in each market (or, more generally, there is no excess demand).

The key feature is that there is *simultaneous* equilibrium across markets. In the full Arrow-Debreu world, prices across history need to be “simultaneously” determined together “outside” the historical process itself. Following the great religious texts, we can place that which is beyond time at some notional beginning of time. Thus we can imagine that at the beginning of time the souls of all the world’s population-to-be assemble in a large building (let us say the Albert Hall). The auctioneer would cry out a long vector of prices covering each commodity over world history. A *tâtonnement* process would occur, until every market was in equilibrium, whether for second-hand animal skins in 2000 BC or microwave ovens in 2000 AD. Once equilibrium had been reached, the final prices would be struck. The souls of the future world would then dissolve from the Albert Hall to return to the unmanifest,

each with a contract. This contract would tell them the prices of each commodity at every time, and their trades. Time would then begin, and souls would become manifest as people and live their allotted lives. History would simply be the working out of the original contract: each day economic man would look at his contract, and carry out his pre-ordained trades (as would economic woman). Economists have competing views as to whether this world would last forever ('infinite horizon' with T infinite), or whether the world would end after some period T . In the latter case, others have argued that perhaps, the world having ended, the whole process would begin again: perhaps the same souls involved in the infinite repetition of the same history, or different souls subject to the same laws. Alas, econometrics remains powerless to adjudicate on this issue. Clearly, the Arrow-Debreu world has little in common with our own world.

A more realistic vision of world history is given by the *temporary equilibrium* approach. This has the advantage that price formation is historically situated. It also has the advantage that each day that economic man wakes up he does not know what is going to happen. The basic vision in its simplest form is to *truncate* the Arrow-Debreu model. At time t , we can differentiate between *spot* markets, which are for goods to be traded in the current period t , and futures markets at which deal with trades to be made in the future, in dates s following t , $s = t...∞$. An extreme form of temporary equilibrium is to assume that the economy is rather like an adolescent punk – *all spots and no future*². Thus at any time there would be a market for each basic commodity, but no markets for future commodities. Suppose an infinitely lived household wishes to sell its 2000 BC sheepskin to buy a 1990 AD microwave. Whereas in the Arrow-Debreu world this can be done directly at the Albert Hall in year 0, in a temporary equilibrium sequence economy the transaction is more difficult. The sheepskin would have to be sold in 2000 BC and exchanged for money (or some other store of value). The money could then be held as an asset until 1990 AD arrives, when it is handed over for the microwave. The reverse transaction is rather more difficult: without a futures market for microwaves, our household would have to persuade the then equivalent of a bank manager to provide it with a bridging loan until the microwave sale in 1990 AD. This should prove no problem with perfect information and perfect foresight. However, if money balances are constrained to be

positive (i.e. no borrowing), then the absence of this futures market might prevent the purchase of the sheepskin in 200 BC.

‘All spots and no future’ is an extreme form of temporary equilibrium. There might be some futures markets allowed (e.g. in financial assets or chosen commodities). In the above example, our household could then trade in these futures markets to finance the sheepskin purchase, e.g. by selling future money for current (i.e. 2000 BC) money. This is in effect borrowing the money to buy the sheepskin.

World history here is a sequence of temporary equilibria. Unlike the Arrow-Debreu world, markets unfold sequentially, rather than all at once. From the household’s point of view, rather than there being one big intertemporal budget covering all of time at once, there is a sequence of budget constraints, one for each point in time (or, with uncertainty, for each state of nature at each point in time). At any one time, all the consumer ‘observes’ are current prices in the markets in which the consumer trades. Whilst the consumer will of course base his current consumption decisions on what he thinks may happen in the future, the absence of futures markets gives him little or no indication of what future prices might be. This contrasts with the Arrow-Debreu story, where all prices for all time are known in the Albert Hall. This of course raises problems of intertemporal coordination. Taking our previous example, if the ancient Briton wishes to exchange his sheepskins for a future microwave, he will have to save (i.e. hold money). This act of saving transmits no direct signal to the microwave manufacturer to invest in order to provide for the eventual demand. This contrasts with the Arrow-Debreu story, where this is all sorted out in the Albert Hall. The absence of futures markets thus poses a decisive problem for coordinating economic activity over time, a problem first highlighted by Keynes (1936) and also highlighted by his subsequent reinterpreters (e.g. Shackle, 1974; Liejunhufvud, 1968; *inter alia*). The first ‘fundamental theorem’ of welfare economics – which states that any competitive equilibrium is Pareto optimal – is only generally valid in the Arrow-Debreu world: in a temporary equilibrium sequence economy the problem of intertemporal coordination is almost insuperable.

Although the Arrow-Debreu and temporary equilibrium approaches seem so different, economic theorists have taken pains to demonstrate that their outcomes need

² For readers not familiar with the late 70’s and early 80’s, one of the themes of the punk generation

not be so different. With perfect foresight, for example, the history of the world (in terms of prices, consumption and so on) will look the same as if history had started in the Albert Hall. Whilst it might be reasonable to assume that infinitely lived households have perfect foresight, it is less appealing in the case of mortal households. The assumption of perfect foresight assumes that each and every agent can work out the equilibrium prices over time. This destroys the appealing notion of ‘decentralization’ by the price mechanism. Rather than each agent simply responding to prices on the basis of his own information, in a perfect foresight sequence economy each agent is required to have full information about the whole economy. In effect, each agent in the economy becomes the auctioneer. Indeed, each agent knows more than Walras’s auctioneer, since he merely groped blindly to equilibrium, whereas our prescient agents calculate the path of prices over history. Of more interest, perhaps, are results which show that a full set of futures markets is not needed to replicate the Arrow-Debreu world: see Marimon (1987) for a survey and discussion of some of the important results.

We have seen the implications of time for Walrasian equilibrium. What of Cournot? The Cournot-Nash equilibrium is best conceived of as a one-shot game. What happens if firms compete over time? We shall examine the issues raised by competition over time to explore the concept of *subgame-perfect* equilibrium; for a fuller exposition, see *Oligopoly theory made simple*. The basic idea of a subgame-perfect equilibrium is that at each point in time agents choose their best responses to each other, given that, in each subsequent period, they will continue to do so. The main use of this equilibrium concept has been to rule out non-credible threats, by which it is meant threats that it is not in the interest of the threatener to carry out (put another way, if the threatened agent called the bluff of the threatener, the threatener would not carry out the threat). The logic of subgame-perfect equilibrium rules out such non-credible threats by requiring agents to act in their best interest for each and subsequent periods. For example, I could threaten to kill you in a most unpleasant manner if you did not send me a £5 postal order. This is not a credible threat since, were you not to comply, it would not be in my interest to kill you (certain readers excepted). Suppose that the Cournot scenario is repeated for T periods. The ‘subgame’ for each period $t = 1, \dots, T$ is simply the remainder of the game from

was that they had “no future”.

period t through to T (for example, in period 5 of a ten-year game, the subgame consists of periods 5 through to 10). At each point in time t , we can imagine each Cournot duopolist choosing his outputs for the remaining periods (i.e. choosing outputs to produce in each period $s = t, t + 1, \dots, T$). If the outputs chosen over the remaining periods by each firm are the best response to the other firm's choice, there is a Nash equilibrium in the period to subgame. Subgame perfection requires that the firms' strategies are a Nash equilibrium in each and every subgame. This rules out non-credible threats, since in effect it requires the firm to choose its strategy optimally at each stage of the game. When we arrive at the period when the non-credible threat needs to be carried out, the firm will not do this because it is not in its interest to do so.

In order to solve for a perfect equilibrium, agents need to 'work backwards'. In order to know what they will do in period $T - 1$, they need to work out what they will do in period T to evaluate the consequences of their choice of strategy in $T - 1$. They then proceed to work out what will happen in period $T - 1$ conditional on the choice of strategies in $T - 2$ and so on. In a perfect equilibrium, agents in effect become super-rational game theorists: in evaluating the consequences of their actions, each agent takes into account that future responses by all agents will be chosen optimally.

This equilibrium notion has been very popular in the last decade. However, it has a serious methodological paradox at its heart. If the players of the game are super-rational and solve the game, then the equilibrium will occur. However, in evaluating the equilibrium, the agents considered the possibility of taking non-equilibrium actions (i.e. actions off the equilibrium path) but, if agents were rational, then they would never take actions off the equilibrium path. If an agent were to move off the equilibrium path, then he could not be 'rational' in the sense required by the equilibrium concept. Indeed, what would his fellow players think of the player who deviated from the equilibrium path? It is as if the rationality of the players binds and constrains them to a certain course of action. Thus the equilibrium property P2 which defines equilibrium is in conflict with the notion of rationality underlying the equilibrium. Equilibrium is defined by comparison with a hypothetical deviation from equilibrium: yet no such deviation is consistent with the rationality of players.

There are of course ways of attempting to resolve this paradox: the players may be rational, but may make mistakes in choosing their actions (their hands might 'tremble'). However, the paradox remains and stems from the same source as the

general problem with disequilibrium analysis that we noted in section 2.2. If agents deviate from equilibrium, then in some sense they could do better: they have made a mistake. A formal notion of economic rationality cannot explain such mistakes: as economists we cannot understand such mistakes.

To close this section on time and equilibrium, it is worth noting that in the case of an intertemporal equilibrium (e.g. the Arrow-Debreu world or a perfect equilibrium) the stability property P3 is redundant. In an intertemporal equilibrium, actions are made consistent across time. The question of how you get to equilibrium is not possible: you are always already there. In the case of a static equilibrium, you can imagine starting in a disequilibrium state and moving towards equilibrium. If equilibrium spreads across time, there is no possibility of a pocket or era of disequilibrium. This comment is still valid in adjustment models: there is an equilibrium adjustment path to a long-run steady state equilibrium.

2.5 Conclusion

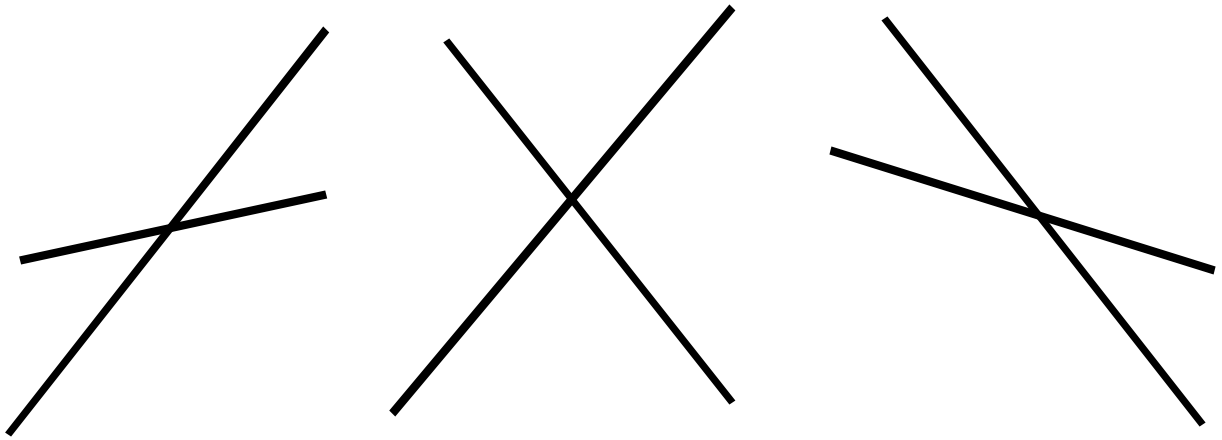
In this chapter, some important equilibria have been looked at, and others have been ignored. What general lessons can be learned from the exercise? I will conclude with personal thoughts.

First, although there is a loosely defined equilibrium method employed in economics, there are many different types of equilibrium which embody different views of the world. It makes little or no sense to talk of ‘the equilibrium concept in economics’. At most there is a family resemblance present.

Second, the equilibrium method plays a crucial role in the process of explanation in economics. Out of equilibrium, actions of agents become inconsistent, or plans cannot be realized, or agents do worse than they could. Mistakes are made. With a formal notion of rationality, economists cannot explain mistakes. In equilibrium, in contrast, the interactions of agents are brought together and made consistent, and in some sense their actions are the right ones to take. We can thus explain things by saying: ‘if everyone does as well as they can, then this happens’. This puts a great constraint on economics. Because economists seek to explain, they seek to expand the equilibrium method to embrace more and more phenomena, real or imagined. Thus what is initially seen as a disequilibrium situation becomes a challenge to economists, who invent new equilibria to cover it (as in the case of ‘disequilibrium’

macroeconomics resting on fix-price equilibria). This paradox is particularly acute in game theory, where equilibrium is defined by hypothetical deviations from equilibrium which should not occur if agents are rational. The answer to the problem might seem obvious: replace a formal model of economic rationality with a substantive model of rationality. If we model the actual decision-making process, we can then explain why it might go wrong. Whilst this may well prove to be the way forward in future, at present it seems an unacceptable alternative to most: to economists there is no obvious model of substantive rationality that is consistent with economics as practised today (which is of course based on formal rationality). Alternatively, the spread of equilibrium may represent an expansion of our explanatory power, an advance of knowledge.

Third, different types of equilibria embody different visions of the world, and with the passing of time economists' perspectives change. Thus what economists view as the paradigmatic equilibrium has varied over time (as it may also vary geographically). Thus, recent years have seen a shift in interest from the Arrow-Debreu world of price-taking agents to a game-theoretic world of large or small agents strategically interacting within and across markets non-cooperatively. It will be interesting to see what comes next.



(a) Keynesian

(b) Marshallian

(c) Cournot

Figure 2:1 The Three Crosses.

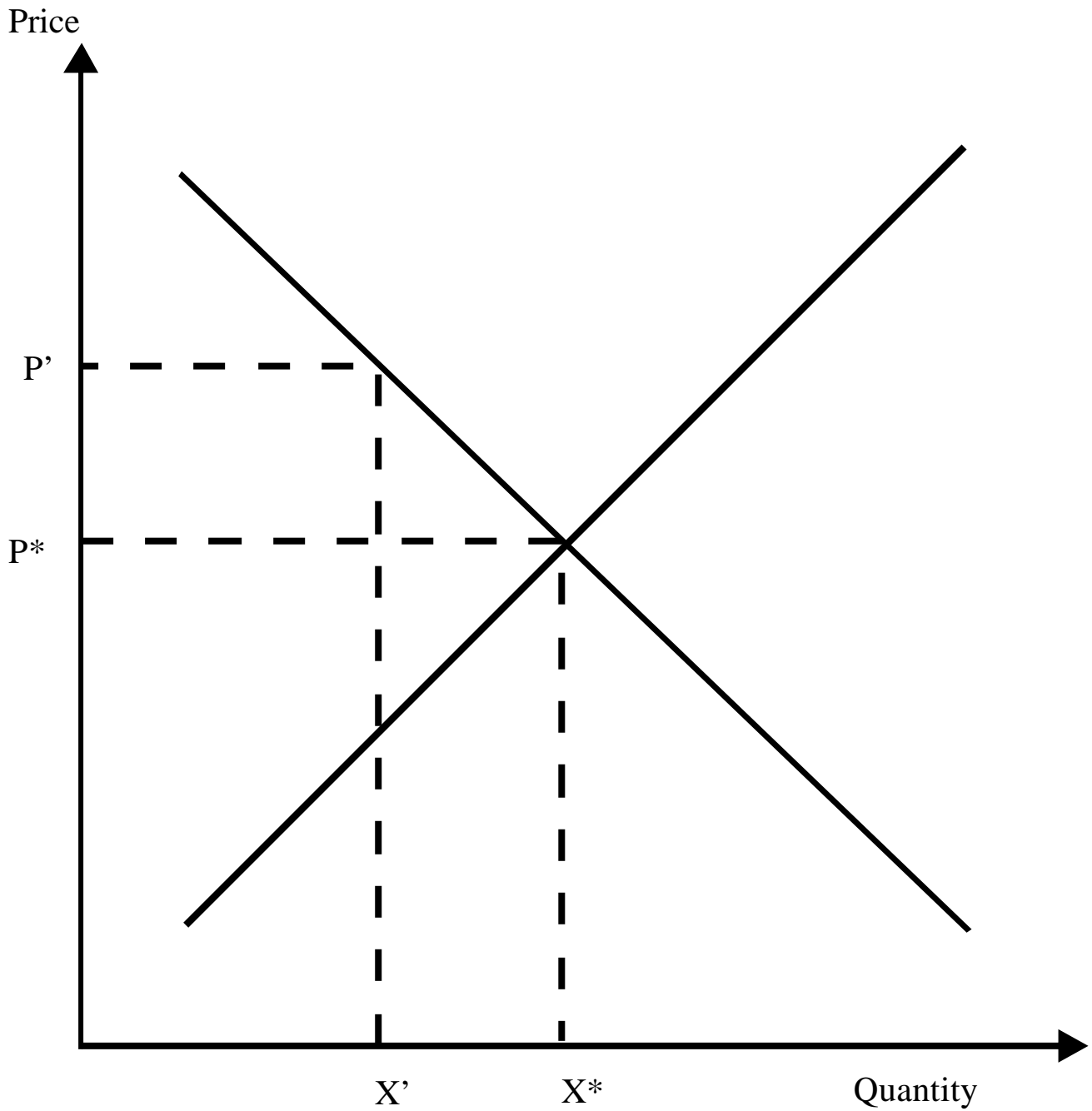


Figure 2.2: Supply and Demand.

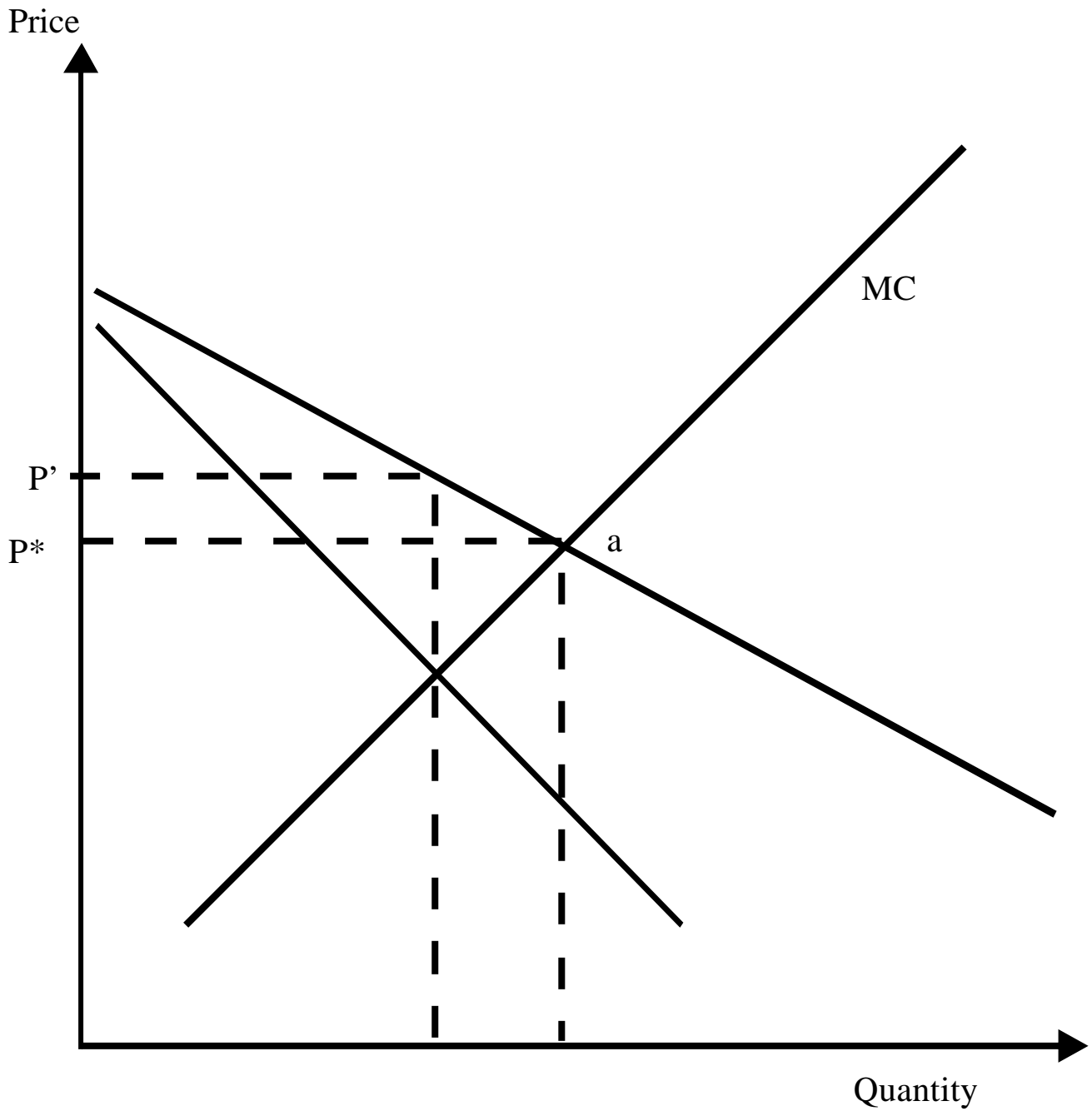


Figure 2.3. The non-existence of competitive equilibrium.

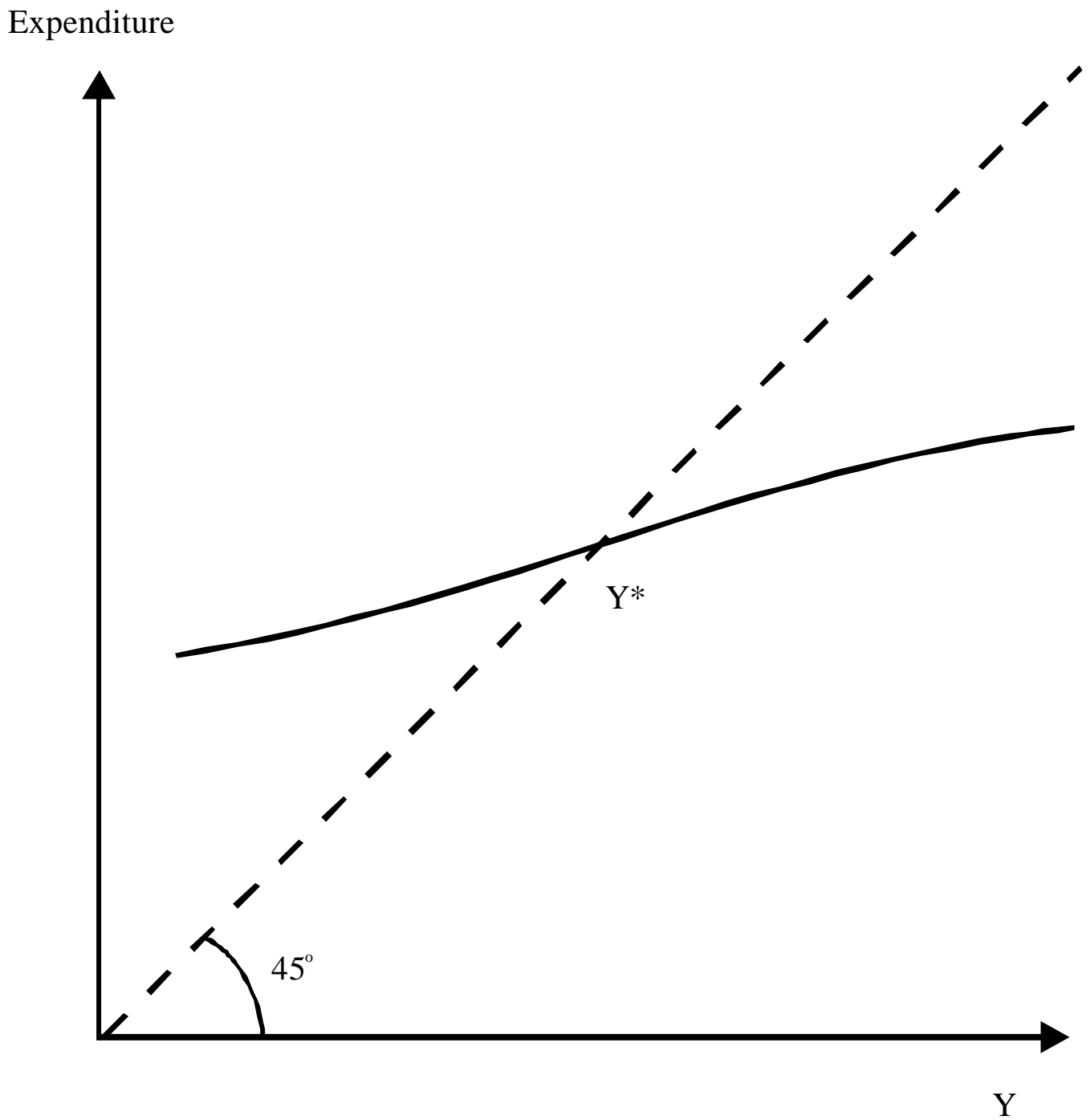


Figure 2.4. Income expenditure equilibrium

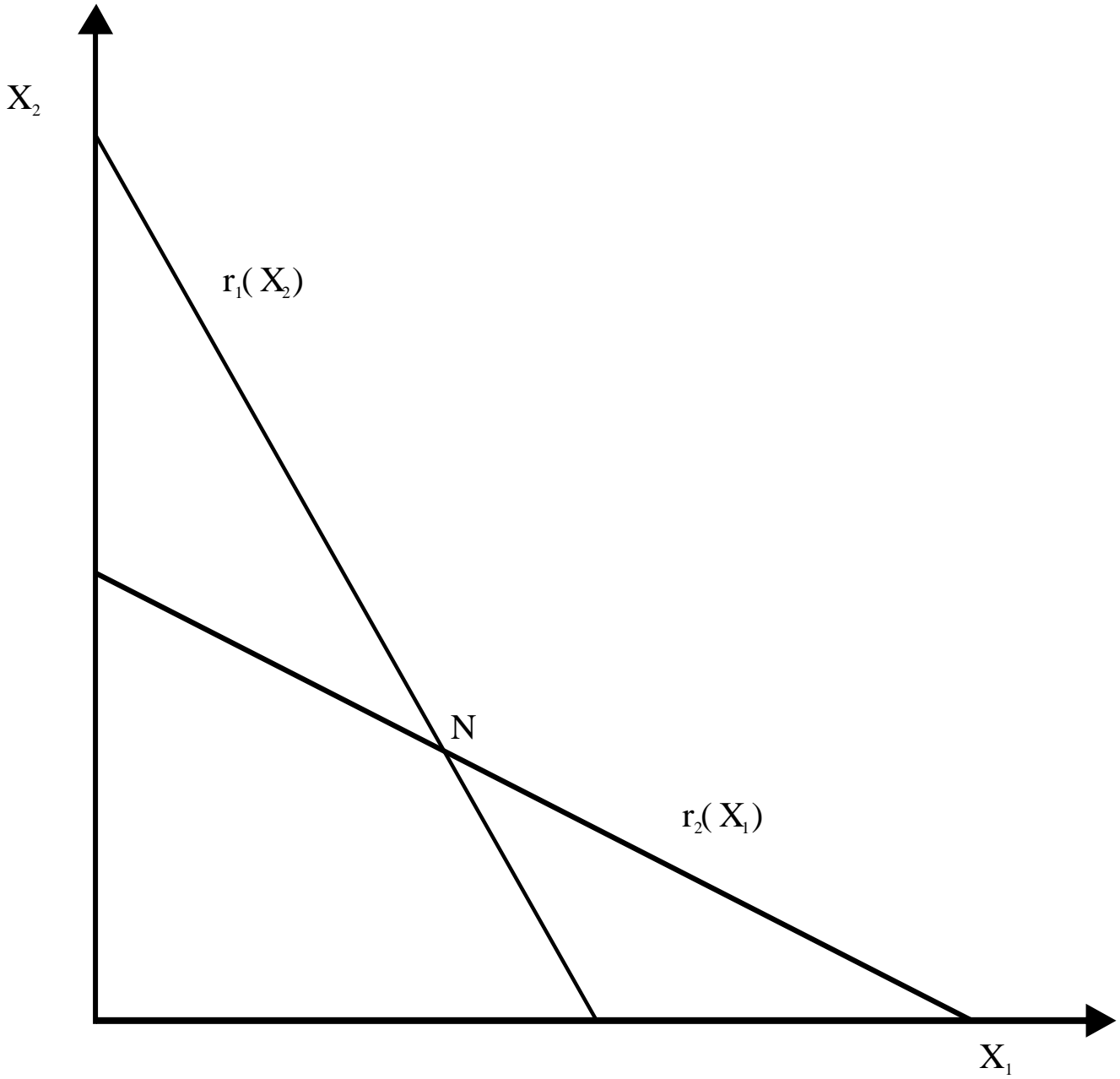


Figure 2.5: Cournot Oligopoly

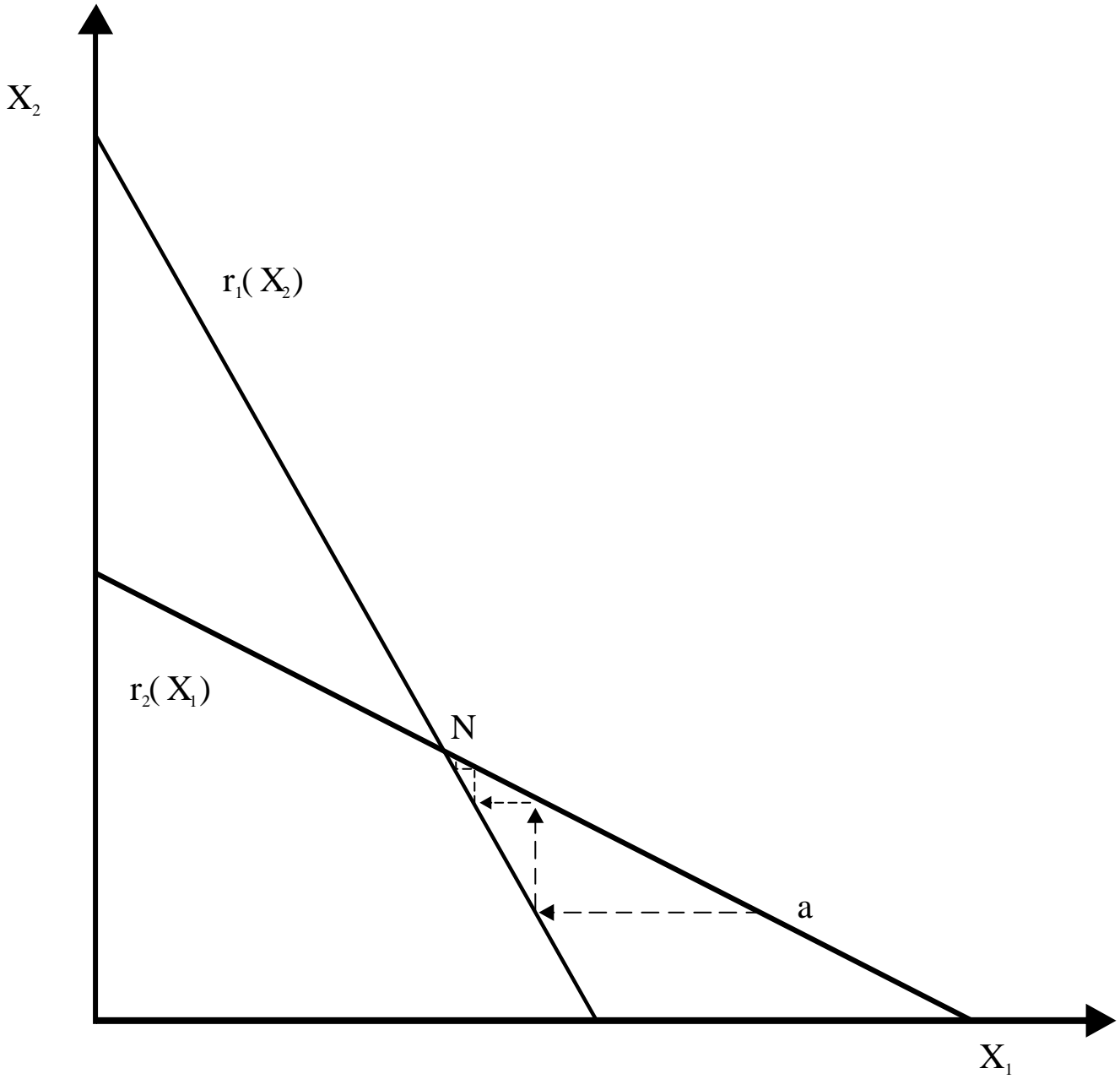


Figure 2.6: Convergence to Cournot Oligopoly

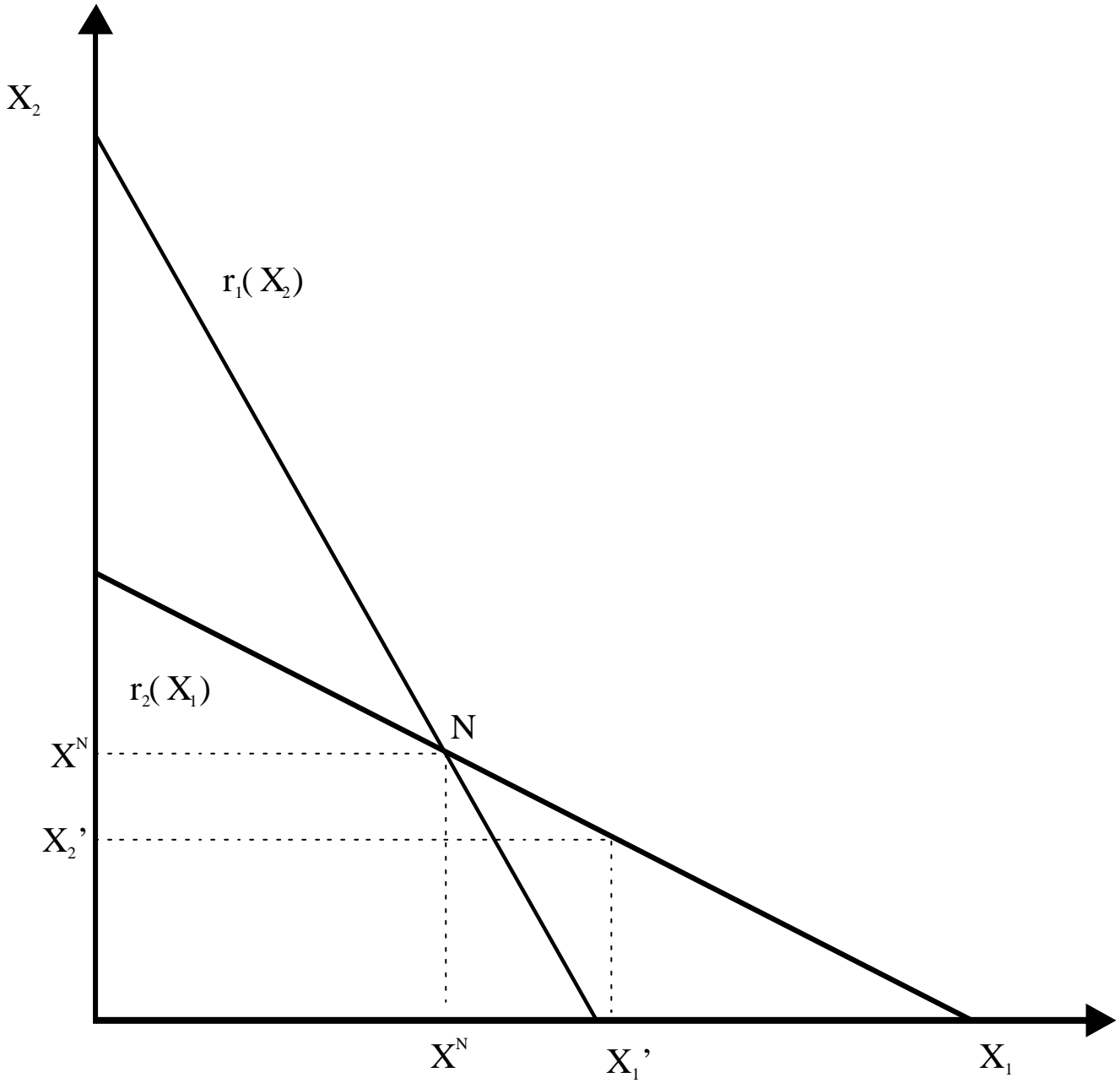


Figure 2.7: Deviations from Equilibrium

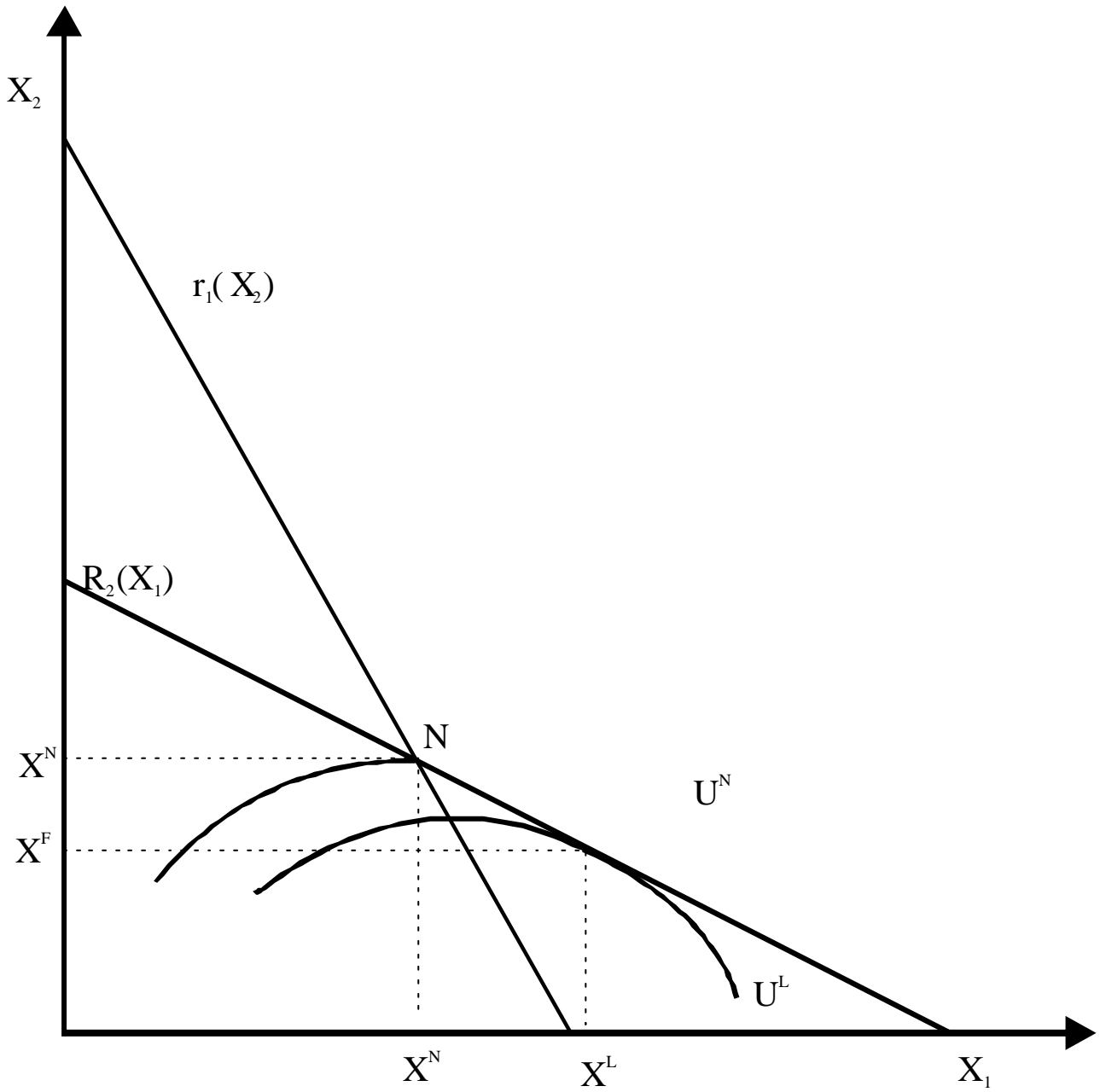


Figure 2.7: Where ignorance is bliss.

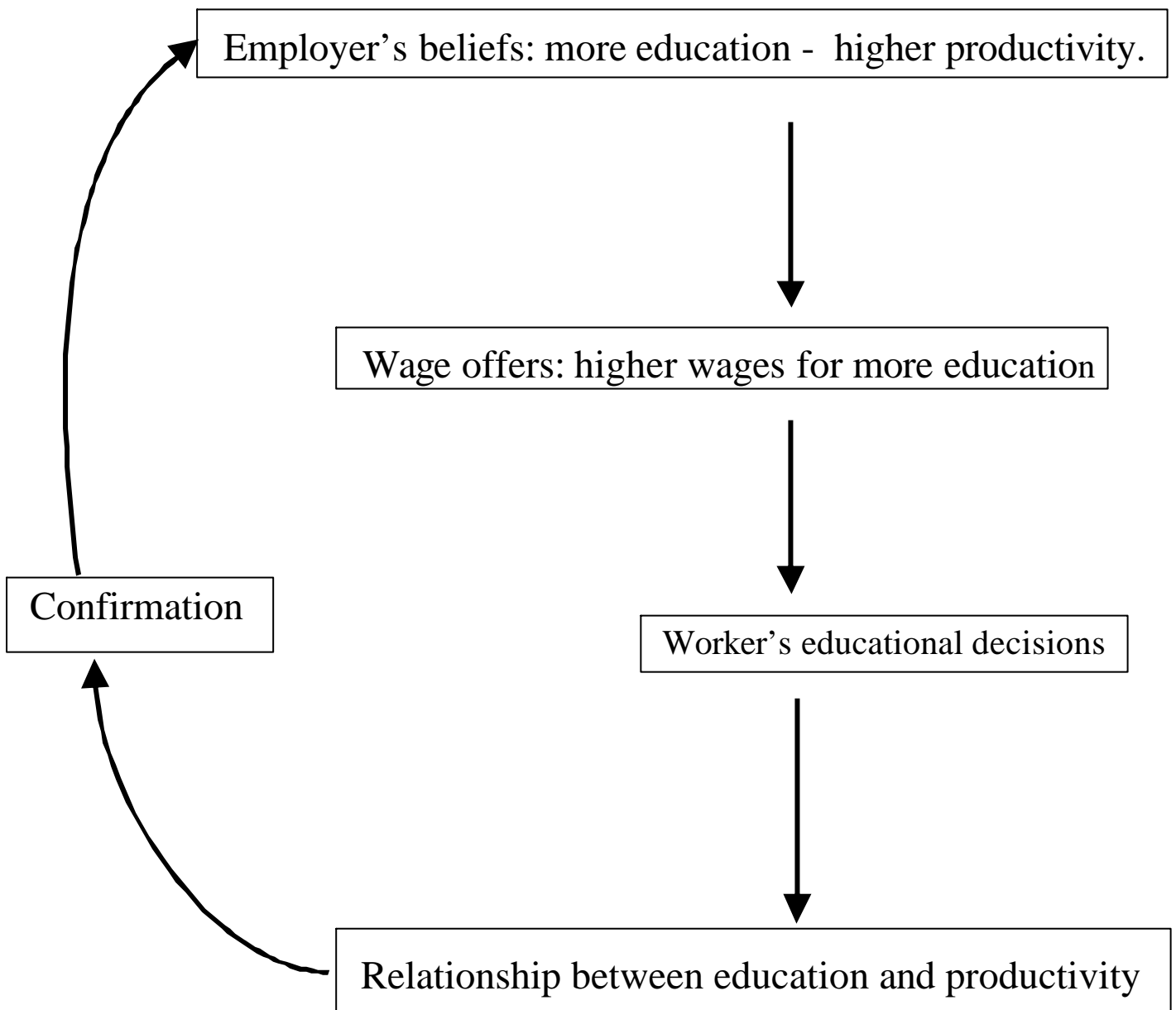


Figure 2.9: A Signalling Equilibrium.

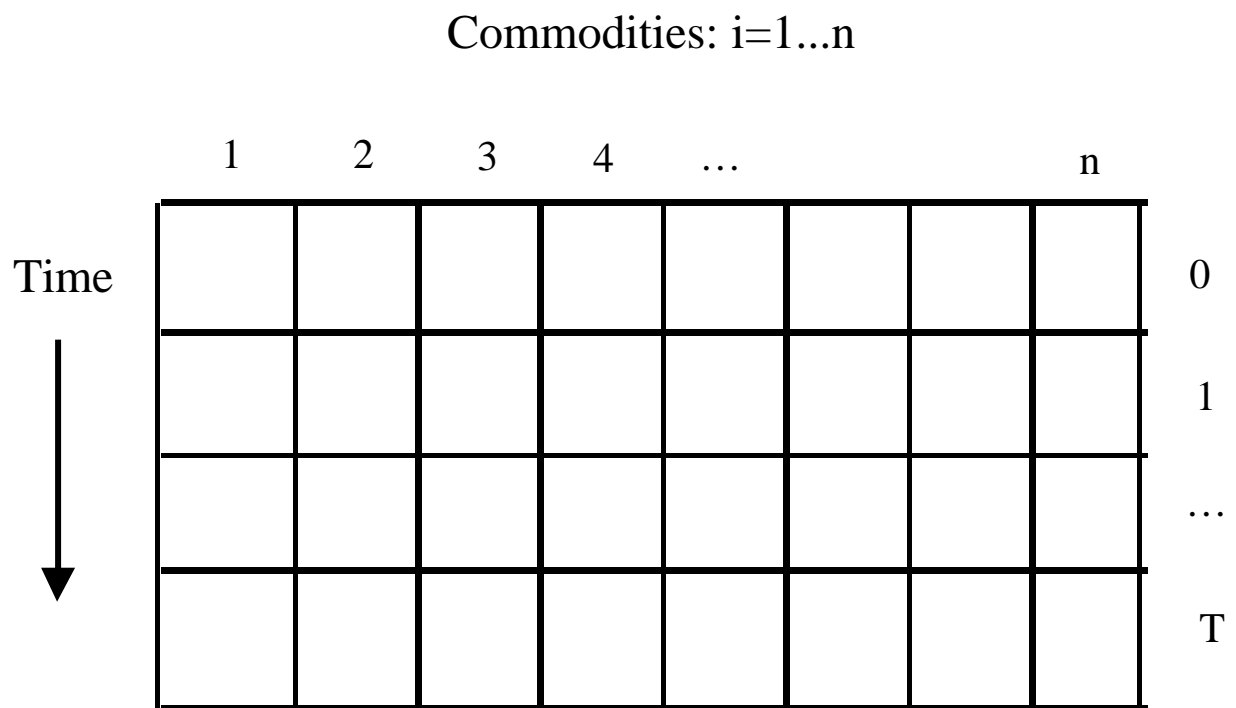


Figure 2.10: The Arrow-Debreu Economy