1 Introduction

1.1 The aims of the book

We had two main aims in writing this book. The first was to explain the prevalence of differentiated goods in a modern market economy and the second was to explore the implications of this for industrial structure and market power. The literature on, and our understanding as economists of, the phenomenon of product differentiation have developed greatly in the last decade and it seemed appropriate at this point to bring this all together and see what general strands could be extracted.

A central question in the theory of imperfect competition concerns the number and kinds of differentiated goods that the market mechanism will produce. One can look at this from both a positive and normative point of view. Our objective is to provide the reader with both points of view in the course of discussing the, by now, fairly substantial literature on product differentiation. Of course, we must also recognise that product differentiation is not the only source of imperfect competition. In its absence, imperfect competition may still arise. This may be for either or both of two reasons: because firms are non-negligible in relation to the size of the market or because consumers or firms possess incomplete information. While we address both of these issues to some extent in this book, they are not the primary focus. For those who want a review of the former, we would refer them to Hart (1985c), while issues of information are dealt with by Butters (1977), Grossman and Shapiro (1984), Rothschild (1973), and Salop and Stiglitz (1977).

Our starting point is the view that what differentiates products are the characteristics that they each possess and so product differentiation involves making a particular firm's product either really or apparently different from that of its rival. However, we need to be a little more specific than this for while chalk and cheese are *different* products, Cheddar and Brie are *differentiated* ones. These examples make it clear that differentiated products are both similar *and* different.

We take it that these differences are grounded in the preferences of consumers. It may be that these are well-founded or they may be engineered by skilful advertising. This is not an issue in this book; we take preferences as data and do not inquire into how they may have come about. We should point out here that, because it provided a framework of reference that enables one to make precise the definition of a particular group of differentiated products, Kelvin Lancaster's (1966) paper in which he provided a framework for modelling choice over characteristics was a fundamental theoretical advance.[1]

However, although Lancaster's paper was an important and seminal one, the discussion of product differentiation did not start with it. The term as such was introduced into the literature by Edwin Chamberlin in 1933 in chapter 4 of his classic *Theory of Monopolistic Competition*. Although the idea was probably used before this, monopolistic competition provided the appropriate context. Chamberlin grappled with two particular aspects. The first was whether differentiation was 'real or fancied'. This led him to focus on things such as trademarks and branding and devote two chapters of the book to the examination of selling costs. The second was how to model these differences. Chamberlin was unhappy with the idea of treating them as 'perfect monopolies, one for each seller', but recognised that it provided a starting point. It was from this that he developed the distinction between the '*dd*' and '*DD*' demand curves that was the novel analytical feature. The former was the (downward-sloping) demand curve faced by the individual firm within a monopolistically competitive group and the latter the group demand curve. The difficulty was that these were taken as primitive ideas in his theory and it was never clear quite how the former was derived from the underlying preferences of consumers. We shall look at this in some detail in chapter 3.

A few years prior to Chamberlin, Harold Hotelling (1929) had published an analysis of competition between firms in a market in which buyers were dispersed geographically and in which firms competed by varying location as well as price. This location model is one we shall deal with in chapter 2. However, the potential application of the Hotelling location model to the analysis of product differentiation was largely ignored until Lancaster's introduction of the notion of characteristics into demand theory. It was this that allowed location and differentiation to be linked.

1.2 Methodology and overview

The approach that we shall follow in this book will be to take as given the nature of consumers' tastes, the specification of technology and a suitable

notion of what would constitute an equilibrium in the particular problem under consideration. In some cases, the latter will be defined in terms of prices, sometimes in terms of characteristics and sometimes in terms of numbers of products.[2] The models that we shall be looking at try to predict the level of product differentiation that would emerge in market equilibrium. This of course is a question of *positive* economics. However, we shall also want to consider the *normative* or *welfare* economics of these market equilibria. An answer to these questions involves comparing the market equilibrium with what we would consider to be the relevant social optimum. In some instances this will be the *first-best* or optimum optimorum, but in others it may make more sense to compare the market equilibrium with a constrained optimum – the so-called *second-best*. This is very much the approach that will be followed in chapters 2 to 6.

A natural question to ask in all of this, and one that would be central in the mind of a policy-maker, is the implications of these market equilibria for industrial structure and market power. Here there are two issues: does product differentiation have an influence on the degree of concentration independent to that of technology and market size and, if the size of the maket is allowed to increase without limit, will the resulting market equilibrium correspond to perfect competition with price being driven down towards marginal cost? More precisely, is the set of limit points of imperfectly competitive market equilibria merely the set of competitive equilibria? This is the theme of chapter 7 where we show that the answer is *no*: there is something fundamental about product differentiation.

While the existence of product differentiation may imply that firms retain some market power, even in large markets with free entry, product differentiation may also impinge in a much more direct and powerful way on market power. This will be so when product differentiation can serve as an entry barrier; in particular, an incumbent monopolist threatened with entry by outsiders into the market.[3] In the product differentiation context this is the analogue of preemptive patenting in R&D. This is an issue that we look at in chapter 8 and it builds on the kind of models that are explored in chapter 2.

For the most part, the analysis we do is carried out in the context of a closed economy. This is a restriction, for competition may be at its most severe at the international level. The new and deeper understanding of product differentiation that we have as a result of recent theoretical advances has important implications for the nature of international competition and is beginning to have an impact on the theory of international trade.[4] This international dimension is an important one for, through trade relations, market size is increased. In chapter 9 we consider

the implications of alternative forms of product differentiation for the theory of international trade using the ideas that we have developed in the course of chapters 3 and 6.

As should be clear from this discussion (and from the title of the book), this is a book about the *core theory* of product differentiation. This inevitably means that a number of topics, important in themselves, but side issues in terms of core theory, had to be left out. An obvious example is *advertising*. This is a topic that has been extensively and accessibly treated elsewhere (Koutsoyiannis, 1984, Schmalensee, 1972). In deciding what we should include in this book, we have taken the view that the most significant theoretical issues in the literature have been the distinction between horizontal and vertical differentiation and the nature of their market equilibria, the advances in the analysis of locational equilibrium and the limit results. We have therefore sought to provide an extensive in-depth treatment of these topics and of some of the more fundamental applications of the theory.

1.3 The formalisation of product differentiation

What differentiates products then is the characteristics that they possess. If we think of each characteristic as being represented by a dimension in some appropriate dimensional space, any product can be thought of as a point in the space spanned by the axes. The relevant set of goods is thus represented by a number of points or vectors in this space.

In explaining the widespread prevalence of differentiated goods, we focus on models in which the choice of product by the firm is endogenous. This obviously makes the question a more complicated one to consider than if we had taken brand positioning to be given, but it clearly seems to be the more relevant framework. Modern corporations take brand positioning seriously and devote considerable resources to establishing and re-establishing their different brands.

Goods can be differentiated in a number of ways. A distinction that we make considerable use of in this book is between goods that are *horizontally* differentiated and those that are *vertically* differentiated.

If all the points in the characteristics space corresponding to the set of goods on offer lie on the same ray vector through the origin, then it is natural to say that a good that is further out along this ray is better: that is to say, is of *higher quality*. If we can find a subset of goods that fits this description, then these goods can be said to be *vertically differentiated* and we could proceed to rank them in terms of some quality index. It would then follow that, if these goods were to be offered for sale at *identical prices*,

every consumer would rank these goods in the same order. On the other hand, if goods cannot be ranked in terms of some quality index, then it seems natural to describe them as *horizontally differentiated*.

As our 'historical' comments above have indicated, the attempts to formalise horizontal product differentiation have followed one of two lines. The first, which has developed out of the Hotelling model, is often referred to as the *address approach*. Tastes are distributed across the characteristics space and individual consumers have their most preferred addresses or locations. In other words preferences are diverse and asymmetric. Each consumer possesses a clear ranking over all available products when they are offered at the *same* price. We can therefore refer to a consumer's 'ideal' or most preferred brand and consider brands close to it in the characteristics space as being better substitutes (and hence, better liked) than brands that are distant. The literature on location models that stemmed from Hotelling has assumed that differentiated products cannot be combined in consumption. This means that each consumer has to choose the one product that best matches his or her preferences after taking into account the costs of acquisition.

If consumers have rather varied tastes, then they will be distributed quite widely across the space whereas, if their tastes are more homogeneous, they are likely to be bunched. Consumers are taken to make rational choices and in this context that means that what they do is to balance the relative prices of goods within the competing group against the relative distances that these are from their ideal specification. Thus the demand for a particular product within the group depends on price, the distribution of consumers and the prices and locations of the rival products. In some cases, this rivalry may be between a few 'nearby' brands in which case the product will have some degree of 'local monopoly' while in other cases the immediately competitive group may be quite large. In general then, the further apart that products are in characteristics space, the less the substitutability between them and the lower the elasticity of demand for any individual product.

The alternative approach which has its roots in Chamberlin is the *non-address* approach. In this case preferences are defined over the set of all possible goods. In a sense one can say here that consumers have *a taste for variety*.[5] A central feature is preference symmetry. The implication of this assumption is that all products are in competition with all others and it means that consumer preferences cannot be used to obtain a ranking over all products when they are offered at identical prices. Thus, while each consumer can have a most preferred brand, symmetry implies an absence of 'neighbouring' brands. Hence, the fact that a consumer is observed to buy brand z provides us with no information about the nature of his

preferences about some brand that is 'near' to z in characteristic space when both are identically priced.

A further subdivision within this approach is that between the case where consumers are assumed to have identical preferences and to purchase all available brands and the case where they are heterogeneous and each consumes a different subset of the available brands. The first of these two cases is the most commonly used and best-known of the models of monopolistic competition: the 'representative consumer' models. A disadvantage of such models is that they don't exactly fit in with certain of what are thought to be relevant stylised facts. Amongst these are, for example, the fact that the range of variants of a product that any consumer buys is in practice relatively limited and such goods are not often *combined* in consumption. Also, different characteristics do seem to matter in different degrees to different consumers and people can often agree on what are and are not close substitutes.

1.4 The incentive to differentiate

If firms in a particular group produce goods which are differentiated, the products of the different firms are imperfect substitutes for each other and, as we hinted earlier, this gives each firm the potential to act as a monopolist in relation to its own product. It is this potential for monopoly profits, due to the fact that it reduces the sensitivity to competitive moves, that provides firms with the basic incentive to differentiate their product. However while differentiation enables a firm to insulate its own market to some degree from the actions of its competitive rivals, the relationship is a symmetric one, and it also makes it harder for the firm to effectively compete in its rivals' own markets. Hence differentiation may well cut it off from a much larger market.

To see the argument, consider a case somewhat similar to the one that we examine in chapter 2. Suppose a situation in which two firms are initially producing the same product and are selling it at the same price. Faced with this rather limited range of products, consumers are likely to make their choice of which firm to buy from in a random manner and each firm will probably end up getting a half share of the available market. Either firm could gain the whole market, for a moment, by under-cutting on price – but it would then lose it all in the resulting price battle as its rival seeks to cut prices in retaliation. In the end price would be pushed down to the zero-profit level. However, if a firm were to vary its product, it would reduce the potential for such unstable competition, and give it some monopoly power in its own localised market. For example, think of shops located along a road. If the store we are considering were to move some

distance away from its rival, this would do some but not lot of good if the vast bulk of consumers were at the original location. However, if at least half of the consumers were in the firm's new market area, the firm would find that it had just as large a market as before, but with the added bonus of a lower elasticity of demand and less competition. It will consequently be able to reap monopoly profit from the move – at least until more firms enter the market.

What this tells us is that increasing the degree of product differentiation gives a firm a far greater degree of monopoly power over an, albeit, restricted portion of the total market. The more varied are consumers' tastes (their own location in this example) and the more uniformly they are distributed, the greater will be the incentive to differentiate – this is, after all, the natural response to the greater taste for variety. However, if the distribution of tastes happens to be concentrated in a particular part of the characteristics space, then we are likely to find that competing firms are producing fairly similar types of good. Their products will be clustered at these high density points and they will be unlikely to differentiate into the thinner parts of the market. This is an effect which is not really taken into account in most of the locational models that we look at in chapter 2 for these assume that tastes are distributed uniformly over a large part of the characteristics space.

As we have said above, for much of this book we implicitly assume single-product firms. The reason for doing so is consistency with the existing literature which assumes single-product firms. However, it should be clear that a large firm may well differentiate its *own* product. We deal to some extent with this in chapter 8 where we explore the issue of the use of *brand proliferation* as an entry-deterring mechanism. There must hardly be a reader who is not aware of the daunting array of breakfast cereals, detergents and toiletries on the average supermarket shelf. Most of these are produced by a handful of firms: two in the case of detergents and toiletries. Indeed, firms will have an incentive to produce multiple brands even when they are monopolists. The reason is that, because consumers do in reality appear to have quite varied tastes, a single brand will appeal to only a small section of the market unless its price is reduced sufficiently to ensure it covers the market. However, the price reduction needed to do this may be one that would leave the firm with little in the way of profit. The advantage of producing more varieties is that this increases the total revenue that can be obtained from selling a given quantity because, on average, consumers find that the available products are more closely matched to what they are looking for and so they will be prepared to pay more for the privilege of having the good. However, the obverse side of the coin is that the quantity produced of any particular variety will be lower

and this will reduce any benefits that are available to the firm from the presence of economies of scale. Thus the optimal differentiation strategy for a multi-product firm will involve balancing the revenue it gains from increased variety against the average cost increase that results from lowering the quantity of any single product produced. Multi-product firms are a missing element in the literature and this is clearly an area where more research is needed in order to further our knowledge about product differentiation.

1.5 Some basic propositions

Perhaps the most notable proposition in the theory of product differentiation is Hotelling's *principle of minimum differentiation.* This was his claim that two firms would locate as close to each other as possible. However, as we shall show in chapter 2, this is a very special result. In the model that Hotelling proposed, it depends on rather restrictive assumptions. We shall see that introducing other elements of imperfect competition, such as incomplete information, can restore the Hotelling result.

Chamberlin's analysis of monopolistic competition provided a theory of product variety in a market economy. He assumed that firms differentiated their products in an optimal way but provided no explicit decision structure for exactly how this happened. Given this, they then behaved like monopolists within the market for their own product equating marginal revenue to marginal cost and so pricing above marginal cost. If the number of firms were small, all would earn positive profits. However, this would attract new entry. Each new entrant would differentiate his or her product from those already available and this would cause the markets of the existing firms, and consequently their profits, to shrink. An equilibrium would be reached when profits fell to zero. The equilibrium number of firms equals the equilibrium number of products and hence measures the extent of product variety.

If profits converge to zero but price remains above marginal cost then this implies that marginal cost is less than average cost. It therefore follows that, in equilibrium, average costs must be falling. If so, we have the conclusion that a monopolistic competition equilibrium must depend on there being fixed costs or scale economies up to some point – for otherwise we could not have an equilibrium on a downward-sloping part of the average cost curve. The equilibrium degree of product variety is greater, the less the effects of scale economies and the less substitutable are group goods for one another – two basic principles in the determination of product variety that carry over to other market forms.

A zero-profit equilibrium will also emerge in the Hotelling locational model when the market is large enough. This has the advantage that location in the market, and hence variety, becomes a decision variable to be explicitly modelled. Each firm will locate far enough from others so as to have some monopoly power and will thus be able to set its price above marginal cost. Positive profits will bring additional firms into the market, reducing market size per firm and thus lowering profits until equilibrium is reached. In the equilibrium, product locations as well as the number of firms will be determind – namely the degree of differentiation as well as the extent of variety.

If the market structure is that of a single multi-product firm, the degree of variety offered will always be less than it would be under monopolistic competition. No clear results are available for an industry composed of a few multi-product firms, although this is a very important real case.

1.6 Product differentiation and welfare

Product differentiation leads to a variety of products being produced. As a result, it is possible that output levels are lower for each product than might be the case if the industry produced a single homogeneous good. If there are scale economies, the firms in the differentiated industry may not be able to take full advantage of these and so there is a potential for resource waste. In free-entry monopolistic competition, the zero-profit equilibrium in the symmetrical case (price = average cost > marginal cost) is such that all firms are producing a lower level of output than that at which costs are minimised. If the product were homogeneous, the free-entry competitive equilibrium would be one in which all firms produced at minimum average cost and so would involve a more efficient allocation of resources. It is this that gives rise to the welfare problem of product differentiation: the possible gains due to the wider range of varieties needs to be balanced against the possible losses due to the inability to take full advantage of the existence of economies of scale.

Product differentiation is also likely to have distributional consequences. If a single variety is replaced by many, there will be gains to those for whom the initial product was not their ideal – this is a gain due to the better matching of products and tastes. Of course, those who preferred the original will lose, possibly on two accounts: the original variety may no longer be available and even if it is, its price will have risen as a result of the smaller scale of its output.

Early welfare analyses of product differentiation were based on the loss that arose from the inability to fully exploit scale economies. They

concluded that the monopolistically competitive market equilibrium inevitably led to a greater degree of product variety than was socially optimal. However, more recent analyses have taken fuller account of the benefits of variety and shown how the reverse relationship may sometimes be true. This is an issue that we shall continually come across in our investigation of product differentiation in the chapters to follow.