The Evolution of Market Structure and Competition

Paul Ormerod (pormerod@volterra.co.uk)

Bridget Rosewell (brosewell@volterra.co.uk**)**

Laurence Smith (lsmith@volterra.co.uk)

Volterra Consulting Ltd

The Old Power Station

121 Mortlake High Street

London SW14 8SN

April 2001

We are grateful to British Telecom for financial support for this research and to Jeremy Turk of BT for comments on an earlier draft

Abstract

We consider in this paper the evolution of a market in which a single product is produced. The specific focus is upon the consequences of new entrants into a market in which, initially, there is a single monopoly supplier.

We set up an agent-based model of firms and consumers, each following particular rules of behaviour for pricing and purchasing. At the outset there is a single monopoly supplier. By construction, the market eventually contains an indeterminate number of producers, each of which supplies the product either at or close to the minimum price which is both technologically feasible and enables a normal rate of profit to be made. In other words, we begin with a monopoly and end with the perfectly competitive market of standard economic theory.

The model extends beyond the comparative statics of conventional theory. In our model, the market evolves over time, and solutions to the model describe the market structure which evolves and emerges from the process of competition.

Judged on the conventional criterion of the distribution of market shares, at any point in time the market structure is, in general, anti-competitive. But by construction, as the outcome on market price shows, the model is highly competitive in any meaningful sense of the word. The market price which emerges, given the number of firms in the market in any particular solution of the model, is in general lower than that implied by the equilibrium Cournot price, substantially so when the number of firms is small. The model further differs from conventional theory in that, across a large number of individual solutions of the model, there is no connection between the market price and the number of firms.

The potential range of outcomes for the model is wide, reflecting the importance of contingency in the process of the evolution of market structure. A certain amount of the variation can be accounted for deterministically, but overall the results contain a substantial stochastic component.

1. Introduction

Conventional approaches developed to examine competition and monopoly generally examine the market outcomes in individual situations and compare them with alternative potential outcomes. Often, the comparison is made between the outcome of a perfectly competitive – desirable – market and the actual situation that is observed. Policy is then directed towards moving the industry under examination closer to the description of the competitive outcome. This approach does not always produce the intended consequences, however, or leaves a suspicion that market power is not being properly characterised. For example, the changes imposed in the UK on brewers' rights to tie their outlets were intended to support the existence of smaller brewers. This is not what occurred. Equally, repeated investigations of supermarkets have failed to identify the exercise of monopoly power – yet public discontent remains.

One possible reason for these apparent failures is that too little attention is paid to the *process* by which the market might change and develop – or has already changed - in the face of structural change. In addition, there is little understanding of the time over which competitive forces operate and over which stable situations emerge.

This paper focuses on the competitive process in one particular example, where it has already been recognised that conventional analysis may be weak. Dynamic, high tech industries are characterised by factors such as falling average costs over a broad range of output, quick and frequent entry and exit with modest capital requirements, economies of scale in consumption and high rates of innovation. Rapid change means that an analysis of process becomes essential. (see, for example, Posner, 2000).

The American authorities have already recognised in principle the challenges to regulatory thinking which such industries pose. For example, Robert Pitofsky, chairman of the US Federal Trade Commission, stated in giving evidence to a Senate committee on 'Mergers and Corporate Consolidation in the New Economy': 'Merger analysis has moved from strict reliance on structure-based presumptions that focused largely on market share

data to a sophisticated analysis that takes account of the dynamic nature of competition in the real world' (Pitofsky, 1998).

We consider in this paper the evolution of a market in an industry in which a single product is produced. The specific focus is upon the consequences of new entrants into a market in which, initially, there is a single monopoly supplier. The illustrative simulations of our model are based upon an outline of the British telecommunications market in which the product is land-line telephone calls. However, it is not our aim to set up a complete model of this particular market.

The model we develop can be applied in its current form to any industry where technical or regulatory change is making entry of new firms possible. Application to a specific industry simply requires calibration of the model to the particular characteristics of the industry. The paper therefore outlines a template which illustrates a general model of market change.

Our model formalises the principle of contestability of markets. We specify a process which determines the timing and number of firms which enter the market, and the prices which they offer. The reaction of existing firms to new entrants is described, as is the behaviour of consumers when confronted with a new set of prices.

The monopoly may have arisen for a variety of reasons. For example, in the twentieth century, in many Western European countries, the physical distribution network of utilities such as telecommunications, gas and electricity was captured by the state, and a single monopoly producer was given control of the physical network. A more market driven example is that of the establishment of VHS as the industry standard in videotape. As VHS sales increased, the production costs of recorders fell as the scale of production rose. Further, as the product became established as the industry standard, economies of scale in consumption arose through more related products becoming available, such as the wider availability of tapes. The computer industry provides another example, with software becoming written increasingly in a particular operating system.

In each of these cases, the monopolist can be thought of as having captured the sales network to consumers. In the case of utilities, this can mean quite literally the physical network over which the product is delivered. More generally, however, this particular concept of a network need not have a physical presence, and refers simply to the connections from a supplier to the consumers. Consumers can only buy from those companies of whose product they are aware. The phrase 'sales network' in this paper means the set of connections from a firm to consumers. By definition, a monopolist has a sales network which connects it to all consumers in the particular market, whether or not the product is supplied across a physical network.

In this paper, we provide an example which starts by assuming that a monopoly has captured the entire sales network – either as a result of state control or for other reasons. We then consider the impact of new entrants to this market. Such new entrants may be made possible by de-regulation and privatisation, but there is also a more general interpretation where such entrants become possible because of new technology. When new entrants come into a market, their ability to sell depends on the customers that they are able to reach, either directly or indirectly. Unless potential customers are able to consider the new offer, they will be unable to take it up. Each new entrant has access to a sales network, which is an important feature of our model. However, the sizes of the networks across new entrants is fixed as a set of parameters in the model, and we do not in this paper introduce potential new economies of scale in consumption.

We set up an agent-based model of firms and consumers, each following particular rules of behaviour for pricing and purchasing. At the outset there is a single monopoly supplier. By construction, the market eventually contains an indeterminate number of producers, each of which supplies the product either at or close to the minimum price which is both technologically feasible and enables a normal rate of profit to be made. In other words, we begin with a monopoly and end with the perfectly competitive market of standard economic theory.

However, the solutions extend beyond the comparative statics of conventional theory. In our model, the market evolves over time, and solutions to the model describe the market structure which evolves and *emerges* from the process of competition.

Section 2 of the paper sets out the theoretical model. The specification is deliberately kept as parsimonious as possible, whilst at the same time maintaining a reasonable degree of realism, in the chosen context of the telecommunications market. The model, though containing only a small number of behavioural rules and parameters, nevertheless generates a potentially complex range of results. Section 3 presents illustrative examples of results. In section 4, we discuss possible extensions to the model, and some implications are discussed in section 5.

2. The theoretical model

2.1 Description and overview

Initially, there is a single monopoly supplier selling a single product to a large number of consumers. The model evolves on a step by step basis, with each step of the model representing a period of calendar time. In the case of the general specification of the model, a market structure is eventually reached in which there is an indeterminate number of suppliers. Each of these supplies the product at a price below that which obtains in the initial conditions, and which is close to the minimum which is technologically feasible. The original incumbent need not be (though almost always is) one of these suppliers.

It is possible to generate special cases of the general specification in which the original monopolist either preserves its monopoly, or is eliminated completely or almost completely.

We specify a process which determines whether firms enter the market in any given period, and, if so, the number which do so. The price which new entrants offer is chosen

at random from within a specified range, but it is always below the original price of the monopoly supplier.

Each new entrant is given potential access to a fixed proportion of the total number of consumers (which, for simplicity, is held constant over all periods). This can be set to be the same for each entrant, and it can be set so that all new firms get access to all consumers. In general, however, we draw the proportion at random, so that each firm has its own unique set of connections to individual consumers.

There are three reasons why new firms in the market do not have potential access (in general) to all consumers, which can obtain either singly or in combination. First, the regulator could impose restrictions so that, for example and purely by way of illustration from the telephone market, a new entrant could be permitted to offer international calls but not domestic ones. Second, the marketing strategy of the firm may be such that not all consumers are aware that the firm is making an offer in the market. In reality, marketing strategies vary widely in effectiveness, and this is reflected in our model. Third, the firm itself may deliberately target only a small percentage of consumers. In the context of British land line phone calls, for example, several firms now specialise in offering cheap calls to India, say, or to the United States.

The fact that a company may have access to a particular consumer does not mean that he or she will necessarily buy from that particular firm. In each step of the model, each consumer reviews his or her choice of supplier. The consumer can only choose from those firms which have potential access to that particular consumer. The original monopolist has potential access, by definition, to all consumers, but in general other firms do not.

The consumer at any point in time is only permitted to buy from a single supplier. This is not always completely realistic, but is a reasonable assumption to make in this initial specification of the model.

He or she identifies the lowest price to which he or she has access and, in a special case of the model, will immediately switch all his or her business to that firm. In general, however, the switch will only be made with a probability which is unique to the particular consumer.

There are several reasons for introducing this probabalistic element into the choice. Although the product offers of the firms are very similar, they are not perfect substitutes, for two reasons. First, the lowest price supplier may specialise in an offer which is not very important to a given consumer. Someone who makes only local calls will not be interested in a firm which provides only cheap international calls. Second, even within the same segment of the market, such as local calls, the product is not completely homogenous in that consumers may have doubts about the reliability of a previously unknown supplier.

There are two other possible reasons why consumers will not in general switch to the lowest price producer. First, there may be costs involved in switching. To take an obvious example, if changing suppliers involved having to change telephone number, for most people the savings on price would have to be considerable to offset the inconvenience involved. Second, consumers may simply exhibit inertia and stay with their existing supplier, perhaps because the savings involved are small.

In each step of the model, then, a process determines how many firms enter the market (a number which can be zero in any given period), and the price at which the entrants offer the product. Each consumer reviews his or her choice of supplier in the light of the available prices.

In the next period, more new firms may enter. Each firm already in the market is in principle able to react to new entrants by reducing its price to the minimum price offered by any firm, including new entrants. Consumers review their decisions given the new prices offered by firms already in the market, and the prices offered by new entrants. Consumers, it should be said, will only switch to firms which offer them a price lower

than their existing supplier. If they have switched in the previous period, say, and their previous (long-standing) supplier now offers them the same low price, they will nevertheless remain with their new supplier.

Firms differ in their ability to reduce prices. More precisely, in this model they differ in the probability with which they can reduce prices in any given period. In principle, any firm is able to adjust immediately to the lowest price which is on offer, but firms differ in the likelihood of so doing. In practice, of course, a firm may not be able to do this at all without, for example, major restructuring of its putty-clay capital stock with its associated level of technology, and so may be driven out of business before it has chance to adapt. However, in this particular specification of the model we can think of firms as differing only in their level of X-efficiency. In practice, differences here, too, may be difficult to eradicate and may persist for long periods of time. But in the first instance, we permit all firms to achieve the most efficient existing level of operation, albeit with differing probabilities of them achieving it in any given period.

We do not specify an explicit process for exit from the market by unsuccessful firms. Once a firm has entered the market, it is able in principle to acquire customers both in the period in which it enters, and in each subsequent period. A firm may enter in a given period, only to find its price to be undercut by competitors, so that it makes zero sales in that period. By reducing its price to the minimum in the next period, it is possible for the firm to then gain some customers, but this number is likely to be small. Clearly, the more periods in which a firm fails to make any sales, the less likely it becomes that it will do so in future periods. More firms may enter the market, and others already in may reduce their prices to the minimum. Firms which fail to make any sales in the period in which they enter the market will, on average, remain in this condition, so their 'exit' takes place at the same time as their entry.

More generally, we do not specify any minimum number of sales which are required for a firm to remain in the market. Once a firm enters the market, it remains there, regardless of its level of sales.

The model evolves on a step by step basis, with new firms entering until a specified total number of new entrants has been achieved. The model continues to be solved period by period after this point, but it does not usually take a large number of such steps before a fixed distribution of sales emerges across those firms which remain.

2.2 A formal statement of the model

The market is populated by n consumers. We assume for simplicity that they each consume an identical amount of the product in any given period. The amount spent per period by each consumer, and hence total sales of the product, may rise over time, but our interest in this paper is on, amongst other things, the market shares of the producers rather than on the total size of the market. So the amount spent by each consumer is the same in any *given* period.

Initially, the market contains a single monopoly supplier, selling the product at a price of p_{mon} (using the subscript 'mon' to indicate the incumbent firm's monopoly price). The model evolves on a step-by-step basis, in which each step is a period of time. We specify a process by which other firms enter the market, both in terms of frequency and in terms of the total numbers entering each period. An obvious way to do this is by a draw from a binomial distribution B(m, p), where we specify the maximum possible number of new entrants in any period m and the mean number of new entrants m*p. We specify (k - 1) new entrants in total, where $k << n$.

Each new entrant comes into the market with a price drawn at random from a uniform distribution on $[p_{min}, p_{mon}]$. The price p_{min} is the lowest possible price at which, after the process of technological innovation is complete, the product can be offered and a normal rate of profit obtained by the most efficient supplier. In each separate solution of the model, it will obviously take differing periods of time before the price $p_{min} + \varepsilon$ is offered.

All n consumers are connected on a network to the initial monopoly supplier. This could be in the case of telecommunications quite literally a physical network, but the use of the word 'network' in this physical sense is too limiting. 'Network' in this context means, more generally, that consumers on the network of firm f_i are both aware of the offer from firm f_i and are willing to consider buying from it.

Each new entrant obtains potential access to a network of consumers. We can specify this to be all n consumers for each firm, or for each firm to have access to a fixed proportion of all consumers, with the precise consumers available to each firm chosen at random. More generally, however, and in the results reported below, each new entrant obtains potential access to a proportion of the total number of consumers drawn at random from a uniform distribution on [ν_{min}, ν_{max}], where $\nu_{min,max} \in [0, 1]$. Once the group of customers to which a firm has potential access has been chosen, it is set up immediately. A further simplification is that the group then remains fixed during all subsequent steps in that particular solution of the model.

In each period, each consumer reviews the prices of the firms to which he or she is connected. The consumer switches all of his or her business to the firm offering the lowest price of these (which may not be the lowest price then on offer to other consumers), subject to the following condition. At the outset, each consumer is allocated a 'switching propensity', s_i, which is drawn at random within [0, 1]. The exact functional form which is used in this paper is set out below in section 3.2. If the customer identifies a price, p_j, which is lower than that of his or her existing supplier, p_i, he or she will then switch to firm j from firm i , with probability s_i.

At the start of the next period, each firm already in the market is given the opportunity to reduce its price to the lowest price which then exists. The ability of the firm to do this depends on the firms flexibility level φ_i. At the outset, each firm is allocated a flexibility level, φ_i, which is drawn at random from a uniform distribution on [φ_{min}, φ_{max}], where $\varphi_{min,\ max} \in [0, 1]$. In each period, each firm switches to the lowest price with probability φ_i.

Consumers then review their choice of suppliers given the revised set of prices from existing suppliers, and given the prices offered by new entrants (if any) in that period.

The model proceeds on a step by step basis, with new firms entering until the pre-specified total number of firms, k, has been reached. The model can then continue to be solved in exactly the same way over subsequent steps, except that there are no more new entrants.

It should be noted that by no means all k firms will have non-zero sales. A firm which is undercut on price and fails to respond may lose all its customers. It is possible for a firm to re-gain some but not all of its previous customers. A simple example will illustrate this point. The model begins with the monopolist, firm 'mon', selling to all consumers. Suppose in the first period, there are two new entrants, firms i and j, with $p_i < p_j < p_{mon}$. A proportion of all consumers will move to firm i and another proportion to firm j, and the remainder will stay with firm mon. In general, neither firm i nor firm j will have access to all consumers, and the sets of consumers to which they are connected will not be identical. At the beginning of the next period, both firm j and firm mon have the opportunity to reduce their price to p_i. Suppose now that firm mon makes this move, but not firm j. Firm mon will regain customers previously lost to firm j because now $p_{mon} < p_j$. But customers lost to firm i will not be regained, because p_{mon} is not lower than, but merely equal to, p_i.

3. The results

3.1 Preliminary remarks

As we noted in the introduction, the model has been kept as parsimonious as is consistent with achieving a reasonable degree of realism in describing how markets with new entrants might evolve. It is, however, complex, and the exploration of its properties is not straightforward. In terms of parameters, for example, we have

- The numbers of new entrants
- the prices at which new firms enter the market
- the flexibility of each firm, measuring its ability to react to lower prices offered by competitors
- the switching propensity of each consumer

The time-scale over which the model converges to a stable distribution of market share at a price close to the minimum possible price can be affected substantially by the choice of the ranges from which these parameters are drawn. Further, the outcome as far as the incumbent monopolist is concerned is affected by its position in the range of flexibility.

An additional complication is the choice of process by which both the timing and numbers of new entrants are determined. This affects not just the time-scale but the actual outcomes of the model, as does the choice of the networks of consumers to which the new entrants are connected.

In this paper, we therefore report on just two separate, realistic scenarios for the timing and the number of new entrants, scenarios A and B respectively:

- a relatively low number of entrants introduced over a reasonably long time-scale, so that in any given period the most likely number of new entrants is zero. Each of these entrants has a potentially high level of connectivity. In other words, the value chosen for ν_{max}, the maximum proportion of total consumers which can be connected to any new entrant, is high
- a large number of entrants over a shorter time-scale. In this scenario, ν_{max} is set at a level considerably below that of the previous scenario.

In constructing these two scenarios under which the model is solved, we have in mind two different types of economic structure. Scenario A, with a small number of potential entrants spread over time, is intended to describe a process in which other large firms

enter the market, carrying out substantial R and D programmes. The firms may be carrying out important brand extensions, or may even be investing in a new and more modern type of physical infrastructure. Scenario B, in contrast, is intended to portray a scenario in which many small firms take the opportunity afforded by de-regulation to enter a market, and to use the existing physical network of the monopolist essentially to re-sell the existing product more efficiently than the incumbent monopolist.

3.2 Time scale and parameterisation of the model

An issue which is frequently neglected in economics is how to translate the concept of time in theoretical models into realistic time scales. Often, the question is avoided altogether by recourse to the method of comparative statics. In other words, the equilibrium solution of a model after a posited change has taken place is compared to the equilibrium situation beforehand, and the issue of time is not addressed. Atkinson (1969) demonstrated some of the potential pitfalls of this neglect, showing, for example, that the implied time scale over which adjustment between equilibrium paths took place in neo-classical models of economic growth could be of the order of 100 years.

Our model evolves on a step by step basis, where each step corresponds to a unit of time. There is no unequivocal, natural definition of the unit of time in the model. But switching by consumers on the basis of price is a central feature of the model. Given that in many utility markets consumers receive bills quarterly, it seems sensible to define a unit of time in the model as being one quarter in real time.

We populate the model with 1,000 consumers. This is a sufficient number such that every eventuality can occur. The use of a greater number of consumers is not required in order to understand the properties of the model.

In scenario A, we introduce 19 new firms over a 10 year period (giving a total potential of 20 firms in the market). The entry pattern is drawn from a binomial distribution B(10, 1/20). So the maximum number of entrants in any quarter is 10 but the mean number of

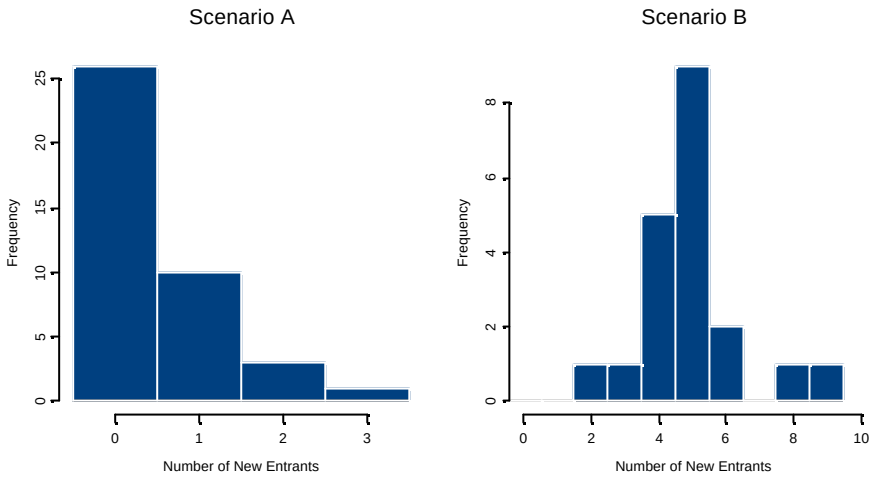
entrants is ½. The draw is constrained so that there are exactly 19 entrants. The proportion of the total number of consumers to which these firms are attached on a network is drawn at random from a uniform distribution on [0, 0.5]. In other words, no new entrant gets potential access to more than 50 per cent of consumers.

In scenario B, we introduce 99 new firms over a 5 year period (giving a total potential of 100 firms in the market). The entry pattern is drawn from a binomial distribution B(10, ½). The maximum number of entrants in any quarter being 10 and the mean being 5. The draw is constrained so that there are exactly 99 entrants. The proportion of the total number of consumers to which these firms are attached on a network is drawn at random from a uniform distribution on [0, 0.1]. In other words, no new entrant gets potential access to more than 10 per cent of consumers.

The percentages of connections are chosen so that the overall level of competition for customers is the same. With 20 companies in the market connected to an average of 25 per cent of customers, each consumer on average has a choice of five providers. Similarly with 100 companies with on average 5 per cent market share, each consumer also has on average a choice of five providers.

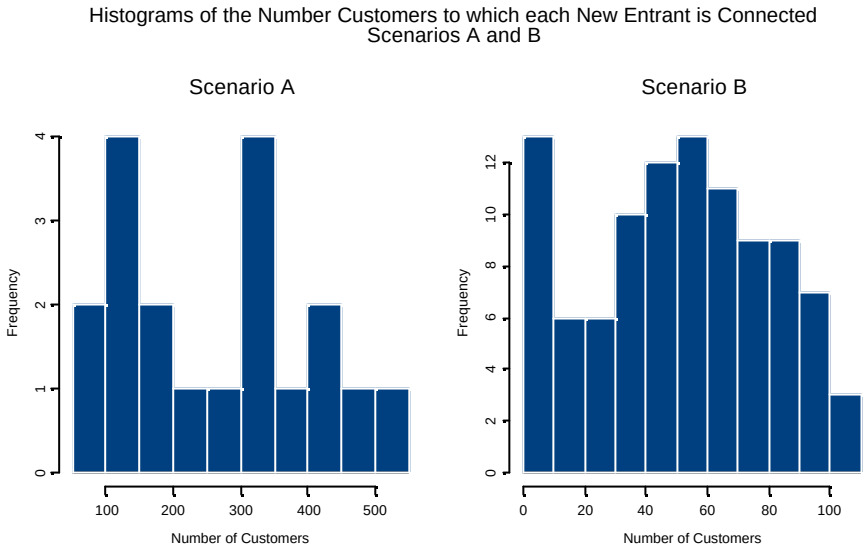
Figures 1a and 1b below set out the histograms of the number of new entrants per period in scenarios A and B respectively, and Figures 2a and 2b plot the histograms of the number of customers to which the new entrants have access.

Histograms of the Number of New Entrants per Period in Scenarios A and B



Figures 1a and 1b

In scenario A, 20 firms are introduced over a period of 10 years, or 40 quarters (40 periods or solution steps of the model). In the majority of these periods, no firm enters. In some 10 periods there is a single new entrant, and only once do three new firms enter in the same period. In scenario B, 99 new firms are introduced over 5 years, or 20 periods. So the frequency with which they enter is quite different to that of scenario A, with at least 2 new firms entering in every single period.



Figures 2a and 2b

The price of the incumbent monopolist is set equal to 1, and the prices of new entrants are drawn at random from a uniform distribution on [0,1]. This does not, of course, mean that the price might literally be zero in reality. The range [0, 1] simply re-scales the range across which prices in any actual situation might vary.

It is worth expanding this point. There are essentially two reasons why prices might fall in this model. First, through competitive pressure as new entrants come into a market previously dominated by a monopolist. Second, as technological innovation over time enables the product to be supplied at prices which were not previously feasible, in the sense that profit could not be made at such prices.

In practice, these two reasons will often be mixed together, but they are conceptually distinct. We might think, for example, of a utility operating with mature technology under a state monopoly. The legal process of de-regulation could permit entrants into this market for the first time, each with access to very similar levels of technology, but operating with varying degrees of efficiency. In these circumstances, prices are reduced

purely by competitive pressures. In contrast, we can imagine a situation in which a company has acquired over time an effective monopoly. A technological innovation is made which enables competitors to enter the market, and continuing investment in research and development enables prices to fall further and further.

Our model does not in its present form distinguish between these two sets of circumstances. In both cases, prices are drawn uniformly from [0, 1]. Yet in practice, prices are likely to fall further in the second of these cases than in the first. The range [0, 1] simply re-scales both these ranges in the model.

This point is important in the interpretation of the results under scenarios A and B. To anticipate, the market price falls more quickly in the range [0, 1] under B than A, for the simple reason that more firms enter the market under scenario B. But this does not necessarily mean that the *actual* fall in price would be more rapid under B than A. As noted above, scenario A is more likely to encompass situations in which other large firms enter the market, carrying out substantial R and D programmes. Scenario B is essentially one of many small firms taking the opportunity to enter a market, and to use the existing physical network essentially to re-sell the existing product more efficiently than the incumbent monopolist. The downward pressures on market price are more likely to include continuing technical innovation under A rather than B. Competitive pressure operates in both scenarios, but in general the absolute fall in price is likely to be higher under A than it is B. In the model, of course, we simply express the range of price under both scenarios as being in the range [0, 1].

The flexibility parameter, φ_i, for each firm is also drawn at random from a uniform distribution on [0,1]. This means that, on average, the incumbent will have a flexibility level of 0.5, exactly in the middle of the distribution. In practice, whether through years of inertia brought about by the previous monopoly position, or because of actions of the regulator, the incumbent may well have a flexibility level which is initially lower than the average. However, we explore here the implications of the incumbent having the average level of flexibility.

The switching parameter, s_i, for each consumer is set up as follows. First, a random draw from a uniform distribution on [0, 1] is made, giving a σ_i for each consumer. The switching parameter s_i is then given by $s_i = 1 - \sqrt{\sigma_i}$. This gives a distribution of the s_i such that most consumers have a fairly low propensity to switch in any given period, and a small number have a high propensity to switch. The probability density function is plotted in Figure 3 below. The mean value of this distribution is 1/3. Supposing, for example, that a consumer with this value of s is confronted on four successive occasions with the choice of switching, the probability that he or she will not do so is just under 0.20. If the s_i were chosen instead simply drawing from a uniform distribution on [0, 1] giving a mean of 1/2, this probability would of course be only 0.0625, a figure which is reached after approximately seven periods with the choice of s_i which we actually use.

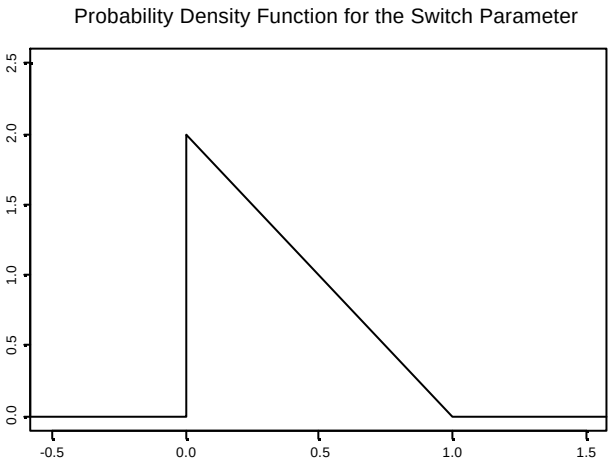


Figure 3

3.3 The results

The probabilistic nature of the model means that its properties must be explored by simulation. Under each scenario, we therefore carry out 1,000 separate solutions of the model, and report the averages over these solutions. Each time, we start from the initial conditions that the monopolist has 100 per cent of the market, selling at $p_{mon} = 1$. The timing and numbers of firms which enter, and their networks to consumers, is fixed

across all the simulations. The scenarios are different from each other in three respects. First, the flexibility of each firm, φ_i, is drawn separately for each scenario. Second, the switching propensity of individuals, s_i, is also drawn separately. Third, the prices at which firms enter the market also differ between the scenarios.

In terms of the output from the simulations, we plot over time the average across the 1,000 simulations in both scenario A and scenario B:

- the market price ie: the prices of firms with non-zero sales weighted by market share
- the market share of the original monopoly supplier
- the number of firms making non-zero sales

We also report, for the most general case of the model, the observed range of these outcomes after periods of 5, 10 and 15 years in the case of scenario A, and after 2 and 5 years for scenario B. The time scale of the evolution of market structure is of course different between the two scenarios because of the different assumptions on the number and timing of new entrants. These periods are chosen to reflect the period of time over which the results settle down.

3.3.1 Special cases

For a number of simplified versions of the model, the results can be deduced in a straightforward manner. The most obvious example is when $s_i = 0 \ \forall \ i$. In other words, there is no switching by consumers. The incumbent monopolist keeps a market share of 100 per cent, and the market price does not change from p_{mon}. However, this same result also obtains when $s_i = \varphi_i = 1 \ \forall \ i$. In other words, when every consumer switches with certainty, but every firm has the highest possible level of flexibility, the monopolist keeps a complete grip on the market. This is because the monopolist immediately matches the price of any new entrant.

In the case where $\varphi_i = 0 \ \forall$ i, in other words where firms do not adjust prices, the outcome is determined entirely by the networks of connections which each entrant draws and the prices at which they enter, and the time-scale is also influenced on the values of the s_i which obtain. In general, in this situation, the market share of the original monopolist declines to zero over time.

The more general case, in which the s_i and φ_i are drawn at random as described, gives much more interesting results. The choice of method by which the s_i are determined is important in two ways. First, it affects the time scale over which the model evolves. This point is obvious. The higher the propensity to switch in any given period, the smaller the number of periods required for the switching process to unfold.

Second, and somewhat more subtly, it affects the eventual market share of the original monopolist. For any given level of φ_{mon}, the flexibility level of the monopolist, the lower the propensity to switch in any given period, the greater the chance the monopolist has to reduce its price to the level of new entrants before consumers make decisions to switch. So, for example, over 1,000 simulations of the model in scenario A with $s_i = 1 \ \forall$ i, the eventual average market share of the original monopolist is 52 per cent. But when $s_i = 1 - \sqrt{\sigma_i}$, as described above, the average is 68 per cent.

The results reported below are for the most general case, in which the s_i and φ_i are drawn at random as described.

3.3.2 Illustrative examples of individual solutions of the model

Each individual solution of the model will be different, because of the probabalistic nature of the model, which is why an exploration of the properties of the model requires a large number of separate solutions. However, before reporting such results, it is useful to examine three individual solutions. First of all, we consider a solution under scenario A, and Figures 4 (a) - (c) below plot the solution over time for the market price, the market

share of the incumbent monopolist, and the number of firms in the market with sales above zero.

Example Plots for Scenario A

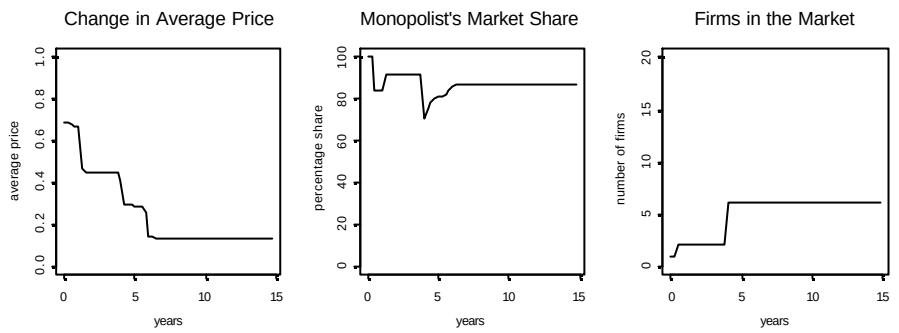


Figure 4

In this particular solution, the monopolist draws a high flexibility (0.73) and so is able to react quickly to new entrants. One new firm enters in quarter 1 and the monopolist immediately adjusts its price to match the price of the new entrant. Thus no customers switch from the monopolist to the new entrant (And so there is still only one firm with positive sales in the market). In quarter 3 a new firm enters at a lower price. The monopolist does not match the new price and so some customers switch from the monopolist to the new provider. In year 2 quarter 2 (ie: period 10), two new providers enter the market. The monopolist matches the new lowest price in the market. It thus regains some of the customers from the second provider who entered and neither of the new entrants in this period get any market share. In the next quarter, the monopolist's competitor adjusts its price to match the monopolist's price and thus the outflow of customers from him to the monopolist finishes.

A similar process is repeated until around the 6 year mark. After year 6, quarter 2 the firm with the lowest price has entered and the monopoly has matched this price and thus the market has stabilised.

In the second example under scenario A, the monopolist draws a very low flexibility (0.005), and the results are set out in Figures 5 a - c below. At no point does the monopolist adjust its price, and thus its market share is continually eroded as providers enter the market at lower prices. The steep fall beginning in year 4, quarter 2 is due to 3 new providers entering at the same time.

Example Plots for Scenario A

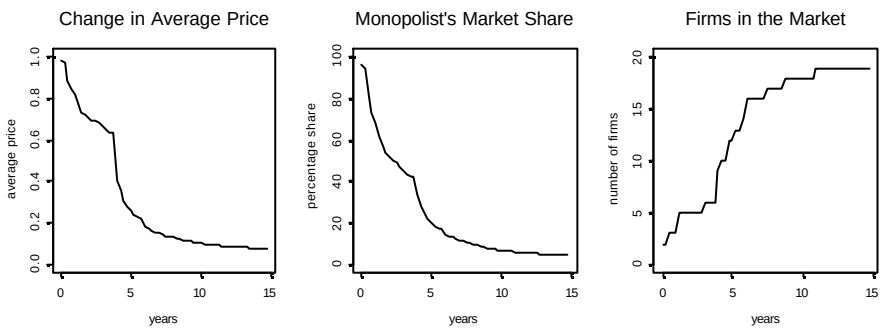


Figure 5

The third example of an individual solution is taken from scenario B, and the results plotted in Figures 6 a - c below.

Example Plots for Scenario B

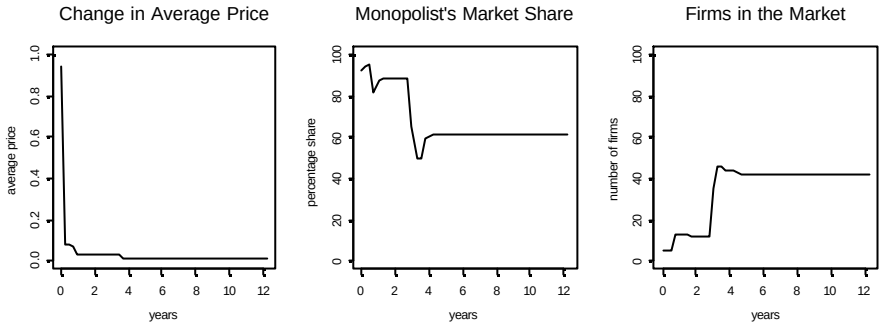


Figure 6

Here the monopolist drew a fairly high flexibility (0.7). Five providers enter in quarter1, each with prices much lower than the monopolists. The monopolist adjusts its price to the lowest price, leading to the very sharp fall in market price. In year 3, the monopolist is losing market share in the first two quarters to new entrants and then adjusts its price to the lowest price and customers abandon the new entrants who do not match the lowest price and return to the monopolist.

3.3.3 The general model, Scenario A

In this scenario, we introduce 19 new firms over a 10 year period (giving a total of 20 firms). The entry pattern is drawn from a binomial distribution B(10, 1/20). The draw is constrained so that there are exactly 19 entrants. The proportion of the total number of consumers to which these firms are attached on a network is drawn at random from a uniform distribution on [0, 0.5]. In other words, no new entrant gets potential access to more than 50 per cent of consumers.

Figures 7a - c plot over time, respectively, the market price (the price of each firm weighted by its market share), the market share of the incumbent monopolist, and the number of firms with non-zero sales. In each case, the average across all 1,000 simulations of the model is plotted.

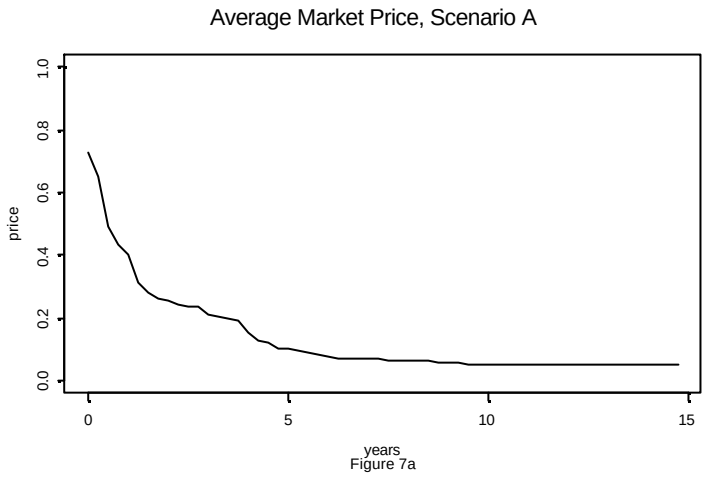


Figure 7a

The market price at the outset is that of the monopolist, which is set to 1, and the minimum possible price which is both technologically feasible and consistent with a normal rate of profit is 0. In the simulations, the market price falls rapidly, so that after 2 years its average value is 0.265. This is reduced still further after 5 years to 0.106, falling to 0.052 after 10 years.

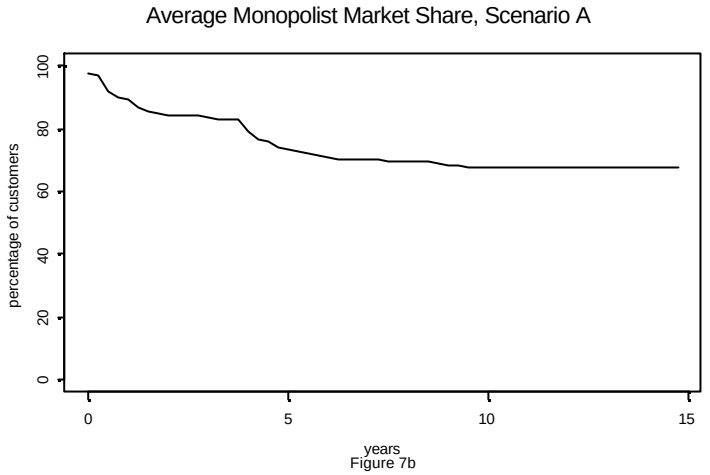


Figure 7b

The market share of the original monopolist remains remarkably high, being some 74 per cent on average after 5 years, and stabilising at just under 68 per cent after 10 years.

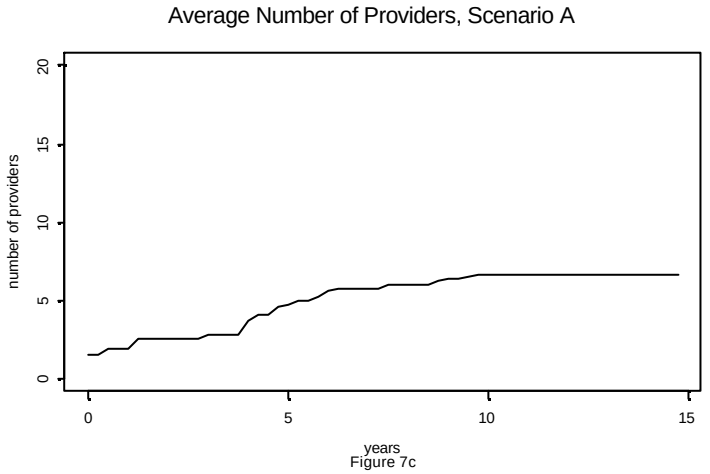


Figure 7c

The chart shows clearly that only a minority of the new entrants survive. At the start of the simulations, there is just one firm, namely the monopolist. The total number rises on

average to 4.6 after 5 years, so only 3.6 new entrants have non-zero sales by then. The figure levels off at around 6.6 after 10 years.

It should be remembered that these are average results for 1000 simulations. Before looking at the range of results, we first review the average outcomes in Scenario B.

3.3.4 The general model, Scenario B

In this scenario, we introduce 99 new firms over a 5 year period. The entry pattern is drawn from a binomial distribution B(10, ½). The draw is constrained so that there are exactly 99 entrants. The proportion of the total number of consumers to which these firms are attached on a network is drawn at random from a uniform distribution on [0, 0.1]. In other words, no new entrant gets potential access to more than 10 per cent of consumers.

Figures 8a - c plot over time, respectively, the market price (the price of each firm weighted by its market share), the market share of the incumbent monopolist, and the number of firms with non-zero sales. In each case, the average across all 1,000 simulations of the model is plotted. The time-scale over which the market structure evolves is obviously much more compressed by construction in this scenario

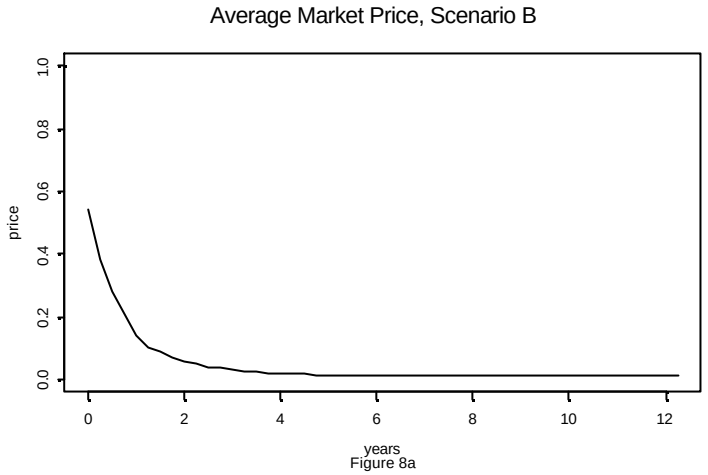


Figure 8a

The market price falls more rapidly in the range [0, 1] than under scenario A, to only 0.07 after 2 years, and 0.016 after 5 years. This is as expected, since the higher the number of entrants and the more rapidly they enter, the greater the probability of very low prices being offered.

However, it is very important in this context to recall the discussions in sections 3.1 and 3.2 above about the economic interpretation of the two scenarios. In both scenarios, the price is expressed as being in the range [0, 1]. But this should be thought of as simply re-scaling the absolute range over which the price might move in any actual situation. It does not necessarily mean that the fall in absolute price would be faster under scenario B than under A. Indeed, the economic circumstances which scenario A is intended to represent is one of large firms entering the market and carrying out substantial research and development expenditure, thereby pushing the market price down through both technological innovation and competitive pressure. Scenario B is intended to represent a set of circumstance in which smaller firms enter following de-regulation, in which further innovation is modest compared to scenario A.

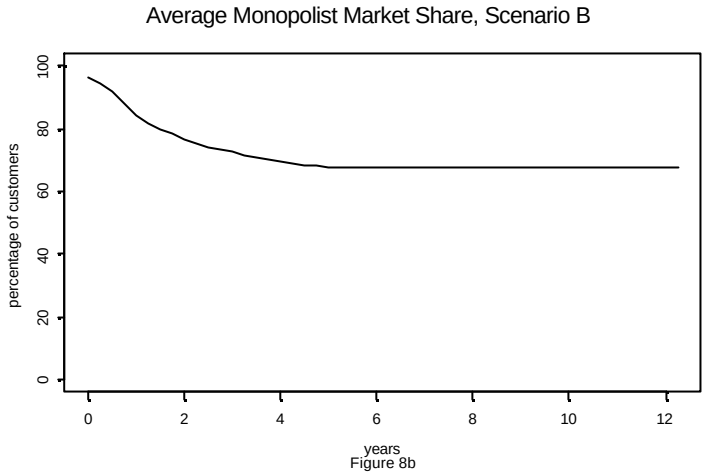


Figure 8b

In contrast to the somewhat different experiences of market price in the two scenarios, the market share of the original monopolist is similar in terms of the level around which it stabilises. In this scenario, on average it is just under 79 per cent after 2 years, and 68 per cent after 5.

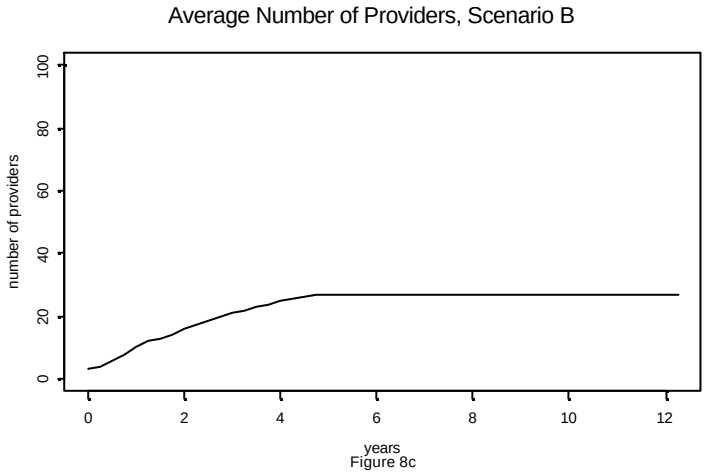


Figure 8c

The number of firms with non-zero sales is obviously much higher under this scenario. After 2 years there are on average 14, and 27 after 5 years.

3.3.5 The range of results

The averages across 1,000 simulations reported above conceal substantial variations in individual simulations. The ranges of results, showing the minimum and maximum values, along with the first and third quartiles, the median and the mean, are shown below.

Table 1a shows that range of market prices obtained under the 2 scenarios at different times, Table 1b the market share of the original monopolist, and Table 1c the number of firms with non-zero sales.

Table 1a Ranges of market prices in 1,000 simulations when the market is fully evolved

Range of Market Prices

	Scenario A			Scenario B	
	5 years	10 years	15 years	2 years	5 years
Min	0.000303	8.88E-05	8.88E-05	5.47E-05	2.63E-05
1st Qu.	0.03408	0.01723	0.01684	0.01023	0.003885
median	0.07914	0.04037	0.039	0.0243	0.008071
mean	0.1056	0.05397	0.05225	0.07086	0.01631
3rd Qu.	0.1593	0.07704	0.0737	0.05012	0.01535
Max	0.4826	0.2992	0.2992	0.4916	0.1782

Table 1b Ranges of market share of the original monopolist in 1,000 simulations when the market is fully evolved

Market Share of Original Monopolist

	Scenario A			Scenario B	
	5 years	10 years	15 years	2 years	5 years
Min	17.8	5.2	3	33.2	9.4
1st Qu.	54.98	42.18	41.72	63.27	43.42
median	81.85	75.95	76.15	86.25	74.8
mean	74.15	67.81	67.59	78.57	68.19
3rd Qu.	100	100	100	100	98.7
Max	100	100	100	100	100

Table 1c Ranges of number of firms with non-zero sales in 1,000 simulations when the market is fully evolved

Number of Firms with non-zero Sales

	Scenario A			Scenario B	
	5 years	10 years	15 years	2 years	5 years
Min	1	1	1	1	1
1st Qu.	1	1	1	1	3
median	3	5	5	11	21
mean	4.641	6.625	6.632	14.06	26.86
3rd Qu.	8	11	11	24	47
Max	12	20	20	39	80

3.3.6 Analysing the results

The tables above show that the experience across the 1,000 individual simulations is, in each scenario, wide. We can, however, account for some of this variation in a systematic way.

3.3.6.1 The eventual market share and flexibility of the incumbent monopolist

Consider, for example, the market share of the incumbent monopolist after 15 years in the case of scenario A, and 12 in scenario B. After these amounts of time, the model has settled down completely to a stable market structure.

In each simulation, the flexibility of the monopolist is drawn at random from a uniform distribution on [0, 1], giving an average value of 0.5. The figure which is drawn for the flexibility of the monopolist is an important determinant of its eventual market share. Also related is the number of firms which remain in the market with non-zero sales.

The correlation between market share and flexibility is 0.76 and 0.75 in scenarios A and B respectively, whilst for between market share and number of firms it is -0.97 and -0.95. Figures 9a and 9b show, for scenarios A and B, the pair-wise scatter plots between these three variables. The three variables are listed on the diagonal set of boxes. The first chart in the top row therefore shows the simple scatter plot of the flexibility of the incumbent on the y-axis against the number of firms with non-zero sales on the x-axis. The chart at the top right is the simple scatter plot of the flexibility of the incumbent on the y-axis against the eventual market share of the incumbent, and so on.

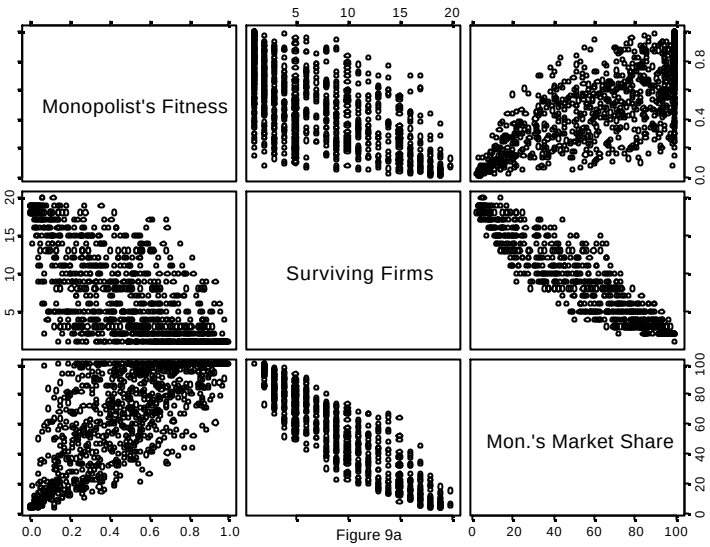


Figure 9a

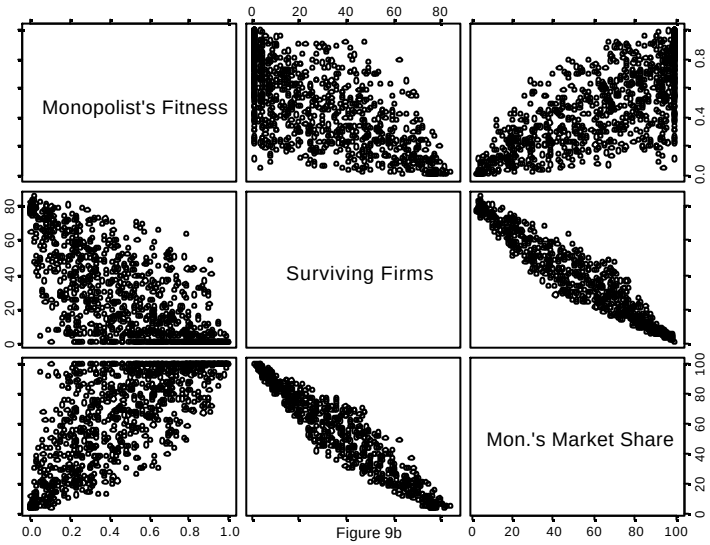


Figure 9b

We can formalise this by, for example, regressing the eventual market share of the monopolist on its level of flexibility. Because market share cannot be higher than 100 per cent, the dependent variable is truncated and therefore we need to use a logit transformation. In other words, the dependent variables is not market share as such, but log[(market share)/ (100 - market share)].

The regressions are very similar in both scenarios, with the estimated coefficients on the flexibility variable, and their associated standard errors, being 4.94 and 0.21 in scenario A and 4.64 and 0.18 in scenario B. The percentage of variation in market share accounted for by the regression is 41 in scenario A and 48 in scenario B. In other words, although the flexibility level of the monopolist goes some way to accounting for the range of outcomes experienced in the two scenarios, there is still a large amount of variation in the eventual market share which is not accounted for.

3.6.3.2 Does the most flexible firm survive?

The concept of flexibility in this model is, of course, that of the ability to respond quickly to competitive pressures. A firm with high flexibility is much more likely to be able to cut its price immediately to match the lowest price on offer than one with a low flexibility.

In this analysis, we exclude the incumbent monopolist because of its special circumstances of access to the network of consumers. We monitor in each solution of the model the firm with the lowest level of flexibility, and observe its eventual market share. The same is done for the firm with the highest level of flexibility. Tables 2a and 2b set out the summary statistics for these factors under the two scenarios.

Table 2a Summary statistics, scenario A : firms with lowest and highest flexibility and the eventual market shares

	min flex	min share	max flex	max share
Min.	0.0000148	0.000000	0.6809	0.00000
1st Qu.	0.0151000	0.000000	0.9293	0.00000
Median	0.0368200	0.000000	0.9635	0.00000
Mean	0.0511100	0.002758	0.9488	0.02845
3rd Qu.	0.0731900	0.000000	0.9857	0.02600
Max.	0.3383000	0.283000	1.0000	0.43400

min flex is the vector of the least flexibility in each simulation; max flex is the vector of the highest flexibility in each simulation; min share is the vector of the end share of the firm with the least flexibility in each simulation; max share is the vector of the end share of the firm with the highest flexibility in each simulation

The difference between 'min share' and 'max share' is clear. For example, the maximum eventual market share of the firm with least flexibility in 1,000 solutions of the model in scenario A is 28.3 per cent, and for the firm with highest flexibility it is 43.4 per cent (recalling that the maximum possible for entrants in scenario A is 50 per cent). A formal Kolmogorov-Smirnov test of the null hypothesis that the vectors of 'min share' and 'max share' have the same distributions is rejected at p = 0.000. So flexibility does play some part in the commercial success of new entrants.

However, a high level of flexibility is no guarantee of success. No fewer than 637 out of the 1,000 firms with maximum flexibility have an eventual market share of zero, compared to 861 of the firms with minimum flexibility. A great deal depends upon the timing of entry. As the model stands, if the most flexible firm enters late in the process, the market price has a high probability of already being very low, so the firm cannot capture any customers. The reaction of the incumbent monopolist is also important,

given its unique network which connects it to all consumers. An early price reduction by the monopolist makes it much more difficult for subsequent entrants to succeed.

Table 2b shows that the results under scenario B are qualitatively very similar (under this scenario the maximum market share of any new entrant is 10 per cent). The null hypothesis the vectors of 'min share' and 'max share' have the same distributions is rejected at p = 0.000 by the formal Kolmogorov-Smirnov test.

Table 2b Summary statistics, scenario B : firms with lowest and highest flexibility and the eventual market shares

	min flex	min share	max flex	max share
Min.	0.0000226	0.000000	0.9269	0.000000
1st Qu.	0.0027870	0.000000	0.9864	0.000000
Median	0.0069840	0.000000	0.9932	0.000000
Mean	0.0102000	0.000119	0.9901	0.005041
3rd Qu.	0.0141800	0.000000	0.9972	0.004000
Max.	0.1109000	0.019000	1.0000	0.078000

min flex is the vector of the least flexibility in each simulation; max flex is the vector of the highest flexibility in each simulation; min share is the vector of the end share of the firm with the least flexibility in each simulation; max share is the vector of the end share of the firm with the highest flexibility in each simulation

3.3.6.3 Market price and market share of the monopolist

We can also consider the outcomes of this approach and compare the results with other ways of thinking about the process of competition. For example, it is often thought that reducing the market share of a monopolist (for example by competition policy) will ensure lower prices. We can examine whether there is any connection here between the

eventual market share of the monopolist and the prevailing market price. Figures 8a and b show, for the two scenarios, the scatter plot between these two variables.

Figure 10a. Scatter plot of eventual market share of monopolist and market price, scenario A

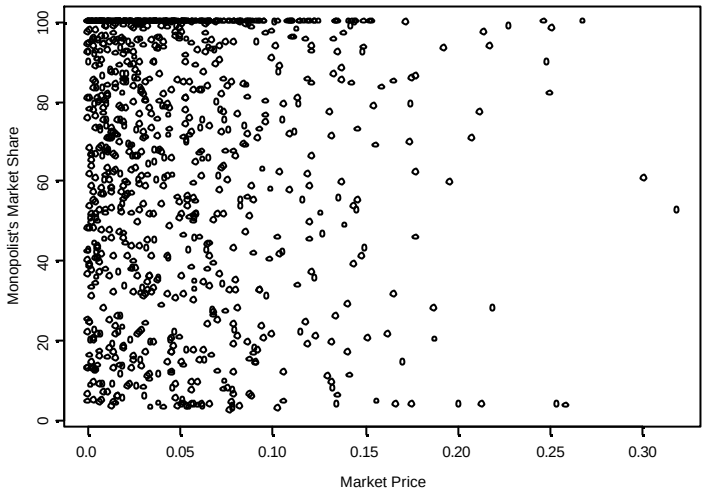
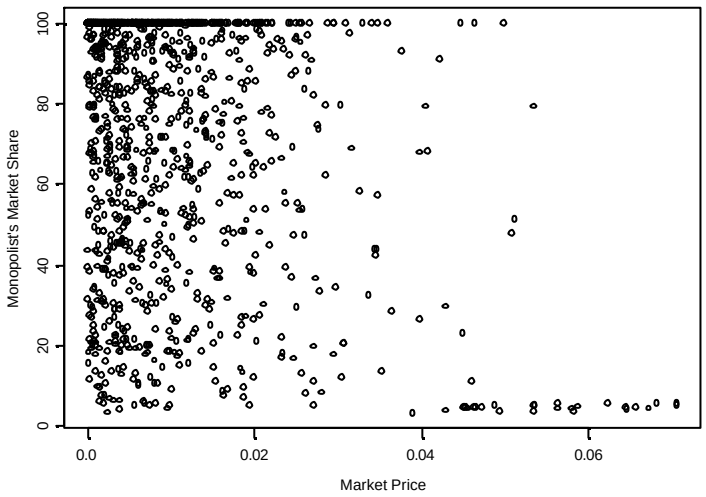


Figure 10b. Scatter plot of eventual market share of monopolist and market price, scenario B



As these plots make clear, there is no connection between these two variables, the formal correlation coefficients being -0.058 and -0.24 under the two scenarios, respectively. An implication of this is that at any point in time, the prevailing market share of the monopolist does not have any implications for the eventual outcome of the market price, and vice versa. A competitive process, in the sense of one which produces a low price, can be associated with any degree of market share for the original incumbent.

3.3.6.4 Market price and the number of selling firms

We can also look at the relationship between the total number of firms and the final market price. Here, standard economic theory implies a relationship between the equilibrium market price and the number of firms in the market. The fewer the number of firms, the more the price will be above the level which just covers both costs and a normal rate of profit.

The Cournot model of oligopolistic markets is used widely. The exact relationship between the equilibrium market price and the number of firms in the market depends upon the particular form of the market demand function which is assumed to exist. As an approximation, however, it is proportional to cost*(1 + 1/N), where N is the number of firms in the market and where cost includes a normal level of profit. As N increases, the price converges on cost.

There are two important differences between our model and the standard economic approach. First, as with the incumbent's market share, there is little or no connection between the market price which eventually obtains and the number of firms in the market. Second, for any given number of firms in the market, the market price in our model is below that of the Cournot model, often substantially so.

Figures 11a and 11b below plot the relationship between the eventual market price and the number of firms in the market under the two scenarios. It is clear that there is little or no connection between the two.

Figure 11a. Scatter plot of eventual market price against number of firms in the market, scenario A

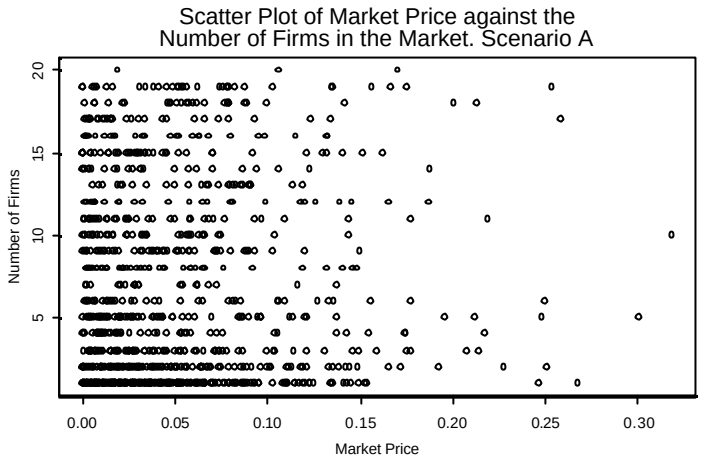
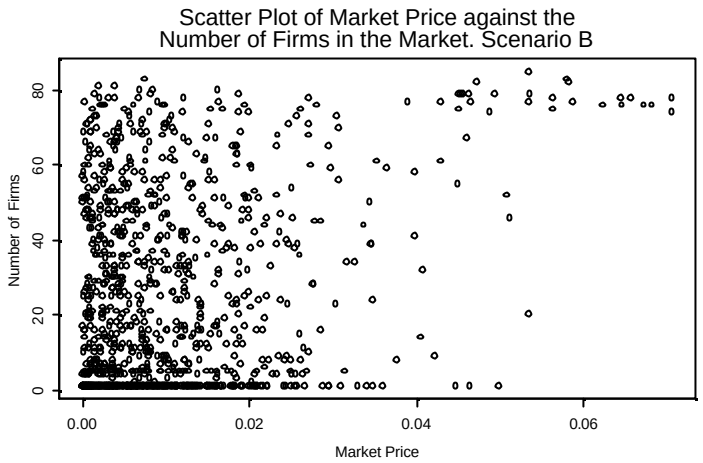


Figure 11b. Scatter plot of eventual market price against number of firms in the market, scenario B



On *average* in our model the more firms there are in the market, the lower is the market price. But there is a strong element of contingency in any particular solution of the model for market price and the number of firms which make non-zero sales.

The lower market price which obtains in our model than in the standard Cournot model can be shown by comparing the average market price which emerges across the 1,000 simulations with the average number of firms in the market.

On average, in scenario A after 5 years there are 4.64 firms in total in the market in the simulations of our model, rising to 5.62 after 10 years. Two widely used illustrations of the Cournot model are with a linear and log-linear market demand function, respectively. With a linear demand function, the mark-up on cost is (1 +1/(N+1)), and with a log-linear one it is (1 + 1/(N-1)). These imply, respectively, a market price after 5 years which is 18 and 27 per cent above cost. After 10 years the figures are 15 and 21 per cent above cost.

A price of zero in our model is defined to be the lowest technologically feasible price which enables costs to be covered, including a normal profit margin. In these simulations, the average market price is 10.6 per cent above this after 5 years, and only 5.4 per cent above after 10 years. In other words, our model implies considerably lower margins than do conventional Cournot solutions. The mere presence of only a small number of firms in the market does not prevent a highly competitive market price from becoming established.

The outcomes in scenario B are closer to those implied by the equilibrium price in the Cournot model. But the average market price is lower in this model than in the Cournot model, even with many entrants. After 2 years, on average there are 14.06 firms in the market and after 5 years 26.86 firms. The Cournot equilibrium price implies margins above cost, for linear and log-linear market demand functions respectively, of 6.6 and 7.7 per cent after 2 years and 3.6 and 3.9 per cent after 5 years. Our results give margins 7.1 and 1.6 per cent above.

3.3.7 Explicit switching costs

In the above results, a consumer switches from firm i to firm j with probability s_i if he or she is on the networks of both firms and if $p_j < p_i$. A simple extension of the model modifies this so that the switch takes place, again with probability s_i, provided that $p_j < p_i - \alpha$, where $\alpha \in [0, 1]$ and is a fixed parameter. In other words, this is a straightforward way of incorporating switching costs explicitly. An alternative interpretation of α is that it expresses brand loyalty to the existing supplier.

This extension to the model by itself makes relatively little difference to the results, except that the eventual market share of the monopolist is in general higher. The likelihood of any given entrant taking customers away in any given period is lower, which helps to preserve the market share of the monopolist. Further, the more steps in the solution of the model take place, the greater is the probability that the monopolist will

reduce its price to that of the minimum and stop further erosion of its market share. With $\alpha > 0$, this reduction is more likely to take place for any specified level of market share than it is for $\alpha = 0$.

Under scenario A, for example, the distribution of outcomes over 1,000 simulations of the model of the eventual market share of the monopolist for very low values of α is not statistically significantly different from 1,000 in which $\alpha = 0$. But even at $\alpha = 0.05$, it is different on a Kolmogorov-Smirnov test at $p = 0.044$.

However, the hypothesis that the distributions of eventual market price are different is not rejected, the relevant pvalue here being 0.770. Even with α as high as 0.5, the pvalue is still only 0.295, so that the null hypothesis that the two distributions are identical is not rejected. The mere inclusion of explicit switching costs does not cause the monopolist, or indeed any other firm, from reacting to lower price offers by reducing its price as well.

To have a more general impact on the properties of the model, explicit switching costs need to be related to another variable, such as the probability of reducing price to the lowest one on offer. The higher are switching costs, the less likely it is that consumers will stop buying from a current supplier, and so the less incentive a firm has to reduce price. This could be extended to make the amount by which price is reduced depend upon switching costs.

4. Potential extensions of the model

The model could of course be calibrated using evidence from a real world situation, but here we are concerned about more general extensions to the content of the model.

In terms of the results, an additional piece of work is to carry out 1,000 simulations of the model for each of a large number of separate draws from the binomial distribution, which

are used to determine the entry patterns of new firms in the market. Similarly, a large number of draws of the networks of consumers, ν_i, which each new entrant accesses ought to be done for the sake of completeness. But neither of these should affect the qualitative properties of the results.

An examination of the consequences of different types of network structure is a much more serious task. At its most basic, this would involve retaining the assumption that entrants get immediate access to a fixed proportion of consumers, with the connections being drawn at random. But the range from which the ν_i are drawn in [0, 1] would be varied from that selected for the results discussed above. This is particularly important in scenario A, where the upper bound on the proportion of consumers to which a firm gets access is currently set at 50 per cent. This could be extended systematically to 60, 70,..., 100 per cent, and the consequences for the eventual market share of the monopolist investigated.

More general still would be to relax the assumption that the size of the network for each firm is fixed, and to allow the network to evolve endogenously within the model. This is particularly important in circumstances in which a relatively small number of large firms enter the market, either for the purposes of brand extension or to invest in an alternative infrastructure. An additional concept of network needs to be introduced, namely the network of the potential spread of information and the spread of decisions by which consumers are connected. We can postulate a range of topologies which connect consumers in the model, and analyse how purchase decisions of the various brands on offer percolate through the network. In part, the decision by a consumer to buy from a particular firm will still depend upon price, but now also upon the decisions of other consumers.

The rules for consumer decisions can be further extended by introducing a process within the model which determines loyalty to an existing brand. In the current version, this is captured by giving each consumer a fixed probability of switching, s_i, so that those with low values of s can be thought of as exhibiting relatively high levels of inertia or loyalty

to their existing supplier. Instead, loyalty could be built up over time as a function of the amount of utility the consumer obtains from each period spent with a supplier. The importance of loyalty in practice will obviously vary depending upon the particular market being analysed.

5 Summary and Implications

This paper has established a model of the competitive process based on a set of rules of behaviour which are both basic and plausible. Firms act in order to enter markets profitably and to increase their market share. Consumers behave rationally, purchasing a new product if it is cheaper.

Our model formalises the principle of contestability of markets. We specify a process which determines the timing and number of firms which enter the market, and the prices which they offer. The reaction of existing firms to new entrants is described, as is the behaviour of consumers when confronted with a new set of prices.

The results of the competitive process, once entry occurs, leads to price in the marketplace which is generally close to the competitive minimum. Individual solutions of the model vary in the speed with which this minimum is approached. In all cases, however, there is a reduction in price and consumers are offered more choice.

In summary, the key results are as follows:

- there is a strong element of contingency in the model, and the evolution of the market over time can differ substantially between individual solutions carried out with the same values of the fixed parameters. However, it is still possible to draw general conclusions about the typical behaviour of the model over large numbers of solutions.

- the market price generally falls from the level set by the initial monopolist to close to the minimum which is both technologically feasible and consistent with a normal margin of profit

- the market price is on average lower than that implied by the Cournot equilibrium, and is usually much lower when there is only a small number of firms in the market

- the monopolist retains, in general, a substantial share of the market

- judged on the conventional criterion of the distribution of market shares, at any point in time the market structure is, in general, anti-competitive. But as the outcome on market price shows, the model is highly competitive in any meaningful sense of the word. The original monopolist retains market share by responding to the competition created by new entrants and by lowering its price accordingly.

The results of this approach to the issue should give regulators and policy makers pause for thought when considering contestable markets. For example, it is not the case that a competitive market (in the sense of having a competitive price), will necessarily have lots of firms, or will have driven down the original incumbent's market share. Further, although market share is often used as an indicator – indeed as a primary indicator – of the presence of monopoly power which may lead to anti-competitive behaviour, these results show that this can be seriously misleading. Finally, the existence of an incumbent by itself does not necessarily tell us much about whether the price is low or high and whether the market is competitive or not.