A Computational Model of Economies of Scale and Market Share Instability

Mariana Mazzucato
Department of Economics
GF, New School for Social Research

Abstract
Replicator dynamics and computer simulation techniques are used to construct a reduced form model which explores negative and positive feedback processes between firm costs and market shares embodied in the dynamics of (dis)economies of scale. After reproducing the standard equilibrium results for decreasing returns to scale (unique equilibrium) and increasing returns to scale (multiple equilibrium), a more dynamic formulation of returns to scale is introduced where scale affects not the direction of costs but the rate of cost reduction. Here we find that negative feedback does not produce self-correcting stabilizing forces in market shares but rather instability and turbulence. Life-cycle phenomena are explored by combining positive and negative feedback in a firm's cost function. The alternating periods of market share stability and instability which emerge from the simulations are compared to empirical regularities in market share patterns.

Keywords: market share, innovation, replicator equations, simulation, industry life-cycle.
JEL Classification: C63 (Computational Techniques), L11(Market Structure: Size Distribution of Firms), O30 (Technological Change).

I am grateful to Duncan Foley for his valuable suggestions and guidance throughout the preparation of the paper. Thanks also go to the organizers and members of the Santa Fe Institute's Summer Workshop on Computational Economics, especially to Yuri Yegorov for stimulating conversations regarding evolutionary cost dynamics. The usual disclaimers apply.
Mailing address: Department of Economics, Graduate Faculty of the New School for Social Research, 65 Fifth Avenue, New York, NY 10003. E-mail address: Mazzucat@is2.nyu.edu.

I. Introduction

The following study uses computer simulation techniques to study the origin and evolution of market concentration and market share instability during the industry life-cycle. Relationships between firm size and innovation are incorporated in a reduced form model which generates a typology of market structures from different positive and negative feedback mechanisms between firm costs[1] and market shares. The typology of market structures which emerge from the simulations are compared to empirical regularities documented by life-cycle studies; market share turbulence during the early stage of an industry's evolution, and market share stability and concentration during the mature stage.

The study differs from previous work on positive and negative feedback between firm costs and market shares (Arthur 1987, 1989, David 1985, 1994) through its focus on the effect of firm size on the *rate of change* of cost reduction rather than on the *direction* of costs. The latter case, which will be referred to here as 'static' economies of scale, has been widely proven to produce unique equilibria of market shares in the case of decreasing returns to scale (negative feedback), and multiple equilibria in the case of increasing returns to scale (positive feedback). This has caused economists, critical of traditional equilibrium analysis, to consider negative feedback processes as uninteresting and to focus attention on the economic implications of positive feedback (Krugman 1979, David 1985, Arthur 1989). A simulation analysis of the former case, centered on the rate of change of cost reduction and referred to here as 'dynamic' economies of scale, finds that *both* positive and negative feedback produce multiple equilibria, and in the case of negative feedback (i.e. when an increase in market share causes the rate of cost reduction to fall) to produce the type of market share instability characteristic of the early period of the industry life-cycle. This suggests that processes of negative feedback should not be ignored by economists interested in disequilibrium dynamics.

The modeling framework is an evolutionary one with focus on the co-evolution of variety between firm characteristics and a selection mechanism winnowing in on the variety. While selection allows only the most fit firms to grow, variety is generated from differences in the ability of small and large firms to innovate and from the interaction of these differences with initial conditions. Replicator equations are especially useful for such a model due to their emphasis on 'distance from

[1] Throughout the paper, the term 'costs' refers to unit costs.

mean' dynamics which views firm evolution not through the representative agent but through the degree to which agents differ from the representative agent; in a world of average behavior there is *no change* (Hofbauer and Sigmund 1988, Metcalfe 1994). The stimulus for using computer simulation to study intra-industry competition is that non-linear behavior is best studied through modeling techniques which do not constrain the degree of feedback amongst the variables. The non-linearity here refers to the possibility that costs affect market shares, which in turn affect costs, together determining the path of industrial evolution; market structure is *endogenous* to the process of innovation. This differs from the one-way causation specified in traditional industrial organization theory between the structure of the market, the conduct of firms , and firm performance (Jacquemin 1987).

The model is composed of 2 x n difference equations; a market share equation and a cost equation for each firm. In each simulation, costs begin randomly distributed while market shares begin equal. The time path of market shares is modeled through a replicator equation which makes the market share of each firm grow proportionally to the distance between its unit cost (fitness) and the weighted average cost in the industry. The time path of costs is modeled through various functional forms which specify the type of feedback (negative, positive, mixed) between costs and shares. Such feedback refers to assumptions on whether it is small or large firms which are more efficient at innovation and how such advantages evolve over time. When large (small) firms have an advantage, i.e. an increase in market share causes the rate of firm cost reduction to increase (decrease), this is referred to as dynamic increasing (decreasing) returns to scale. The life-cycle phenomenon is explored by assuming that such relationships vary during different periods of an industry's evolution.

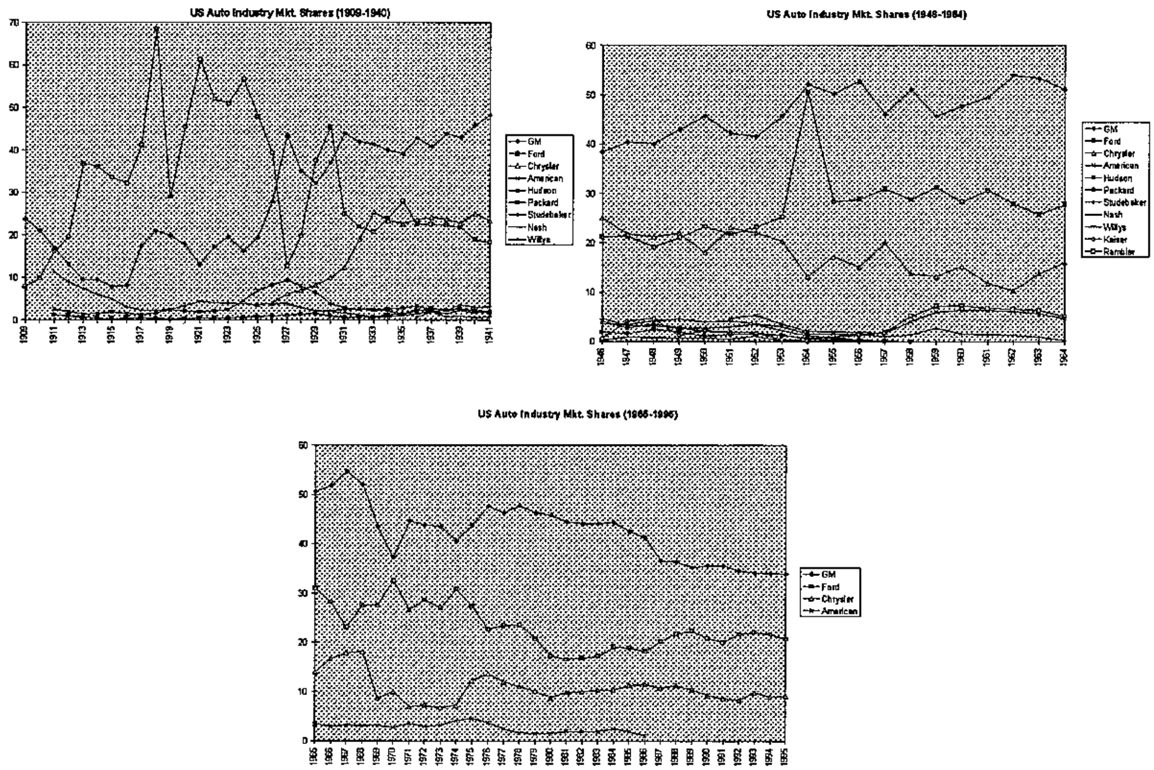
The goal of the simulations is to catalogue the different types of market share patterns which emerge from variations in empirically relevant parameters, and to then compare these patterns to empirical regularities in industrial dynamics such as the skewed size distribution of firms and periodic market share instability. These parameters include the variance of the initial distribution of costs (i.e. the degree to which firms' fitness characteristics differ from each other when they begin competing), the industry-specific average speed of cost reduction (dependent on the speed of diffusion in the industry, the tacitness of the knowledge-base, etc.), the time period in the industry life-cycle in which a change in feedback regime occurs, and of course the functional form of the cost equation which describes the type of dynamic returns to scale being investigated. It is concluded, as in various empirical studies, that the question to ask is not 'Which size firms are more efficient at innovation?' or 'What type of market structure emerges from innovative activity?' but rather 'Under *what conditions* are small and large firms better innovators?' and 'Under *what conditions* do concentrated markets emerge?' (Acs and Audretch 1987, Cohen and Levin 1989). For example it is found that in the case of dynamic decreasing returns to scale, if the industry-specific rate of cost reduction is very slow (inertia) a stable and relatively concentrated market emerges, while if the rate of cost reduction is intermediate a turbulent more competitive market structure emerges. This result is directly related to empirical studies, explained below, which find a connection between the intensity of innovation and market structure (Dosi 1984, Pavitt 1984, Cohen and Levin 1989).

While the reduced form nature of the model implies that it cannot capture the specificities of one particular industry or firm, its purpose is to capture general properties of a wide group of industries which are not easily revealed in very detailed models. Understanding these general properties 'clears the ground' for more specific assumptions to be explored regarding demand, entry, price etc. The abstraction from details allows the analysis to focus on the relatively unexplored disequilibrium dynamics which emerge from negative feedback, and to posit a very clear difference between firm-specific factors and industry-specific factors which affect market structure. One element which is purposely omitted from the analysis is the effect of stochastic shocks on market structure. Such shocks are fundamental to capture the random and uncertain elements of innovation and the role of early 'small' events. For example, Henry Ford's megalomaniac personality surely had a strong effect on the evolution of Ford Motor Company's market share, independent of cost dynamics. Yet in order to uncover the *non-intuitive* properties of the deterministic dynamics, the treatment of randomness is left to a subsequent paper which will focus solely on how the degree to which random shocks influence the deterministic dynamics differs in various stages of the life-cycle (i.e. in different regimes of innovation).

II. Empirical Regularities and the Measurement of Competition

The study is motivated by two regularities widely documented in industrial dynamics : a) the 'skewed-size distribution' of firms found to exist across many industries (Simon and Bonini 1958, Hart and Prais 1965, Ijiri and Simon 1977), and b) the alternating periods of market share stability and instability during an industry's life-cycle (Gort 1963, Hymer and Pashigan 1962, Klein 1977). For example, in Fig. 1a below we see that in the US auto industry there has been a prevalence of periods of market concentration accompanied by alternating periods of market share instability between US producers:

Figure 1a Market shares of US automobile producers [2]:



It can be seen that while the early period (until the late 1930's) was characterized by market share instability and changing levels of concentration, the later period (from the 1940's onward) was characterized by market concentration and stability. The large change in the number of firms, hence also in the degree of concentration, is noted clearly below in Figure 2 (p.9). Similar patterns have been found in other industries such as television set (Datta 1971), aircraft (Phillips 1971), tire (French 1990, Klepper 1996), semiconductor (Gruber 1994). These patterns suggest that there are strong underlying factors generating industry concentration and periodic market share stability.

Explanations regarding the cause of such asymmetry (concentration) and turbulence have focused on two main processes: 1) the optimal allocation of scarce factors of production, and/or b) the

[2] Sources for Figure 1: 1911-1937: Federal Trade Commission (1939, p. 29, 682-683, 715, 749-750, 812) 1940-1995: Wards Automotive Yearbook

random nature of firm growth patterns. The stochastic approach arose principally due to the dissatisfaction among industrial economists with static cost curve analysis which seemed incapable of explaining the role of technology-specific and firm-specific factors determining long-run patterns. This approach typically assumes the size of a firm at time $t+1$ to be a function of its size at time t subject to random variation. The stochastic process most studied in this regard is *Gibrat's Law of Proportionate Growth,* which states that firm growth rates are i.i.d. random variables independent of firm size (Gibrat 1931, Kalecki 1945, Simon and Bonini 1958, Ijiri and Simon 1977). The principle result in such models is that although firms might begin ex-ante with equal growth prospects, random events soon cause firms to diverge in size and a skewed size distribution to emerge (log-normal).

Some problems with the stochastic approach are that 1) it makes no mention of the periodic *instability* in market shares which has been found to exist simultaneously with the skewness in size[3], and 2) that it makes market structure solely a result of random factors, which as is claimed by Geroski et al. (1996): "*may be more an artifact of the models than of the data itself*". The reliance on randomness ignores empirical case studies which refer to specific economic mechanisms, concerning firm size and innovation, which might generate such patterns (Klepper 1996, Audretch 1995). The paper takes an alternative approach to modeling market share dynamics. It emphasizes both structural industry conditions, embodied in cost characteristics, as well as firm-specific and industry-specific variables like the initial random distribution of costs. The structural dynamic described in the cost curves is taken from the industry 'life-cycle' literature where it is argued that industries are characterized by alternating periods of positive and negative feedback between firm size and efficiency (Abernathy and Utterback 1978, Audretch 1995).

For each different cost scenario the level of instability and concentration is plotted. The indices used here to measure competition come directly from the literature concerned with the empirical regularities described above. The primary criticisms made to the traditional concentration ratio and

[3] An exception to this critique is found in Winter and Rothblum (1985), where the simultaneity of a skewed size distribution of firms with market share fluctuations is addressed in a stochastic growth model which follows Gibrat's assumptions stated above. They show that although there is a tendency toward increasing market concentration, sample paths for individual firms fluctuate from very high market shares to very low market shares (from near monopoly to near extinction). They claim that the reason that there is so much emphasis on concentration rather than on fluctuation is that the latter is very slow and hence unrecognizable in the short period. Their model, however, is still susceptible to the second critique since it uses only random factors to explain the regularities of concentration and turbulence.

herfindahl index are that they do not capture the way that concentration and variety *change* over time. The inadequacy of using (only) traditional concentration indices to measure competition can be seen in Figure 1; in certain phases of the history of the automobile industry (ex. 1923-1941) there was much turbulence in firm market shares, indicating the presence of competition, while there was also a relatively high concentration ratio, indicating the opposite. Gort (1963) claims that the strongest argument against the traditional indices is the empirical evidence of turbulence amongst market shares in many industries. He states:

> " *One of the chief objections to 'concentration ratios' as descriptions of market structure is that high ratios may be consistent with considerable instability in the market shares of individual firms. In judging the intensity of competition in an industry, the ability of leading firms to maintain their relative position in a market is probably more significant than the extent of concentration at a single point in time.* " (p. 51)

When measuring the degree of market share instability below, the index devised by Hymer and Pashigan (1962) is used. This index, devised to specifically confront the above problems, is defined as: $I_j = \sum_{i=1}^{n} [|s_{it} - s_{i,t-1}|]$, where s_{ij} = the market share of firm *i* in industry *j*.

The larger is the value of I, the more unstable are market shares in the industry, indicating the presence of competition. To measure concentration, the 'relative' entropy index (i.e. relative to its maximum value) is used. This is defined as: $E = \frac{-\sum_{i=1}^{n} s_i \log s_i}{1/n}$ where n is the number of firms in the industry . When $E = (1/n)/(1/n) = 1$, relative entropy is at its maximum value which means that the n firms' market shares are equal, or in Shannon & Weaver's (1949) terminology, that the "information" content of the industry is at its highest. Since each firm begins with a market share equal to $1/n$, E is always equal to 1 at $t=0$. Jacquemin (1979) and others have shown this index to approximate the inverse of the Herfindahl index.

III. Firm Size and Innovation

The common usage of the term (dis)economies of scale, which is called here 'static' (dis)economies of scale, refers to the (rise) fall in unit costs which arises from an increase in quantity produced. Diseconomies of scale pose a limit to the size of the firm by punishing large size with rising unit costs. It is this aspect of the neoclassical long-run average cost curve that allows market structure to be easily predicted. Arthur states:

"An example is the competition between water and coal to generate electricity. As hydroelectric plants take more of the market, engineers must exploit more costly dam sites, thereby increasing the chance that coal-fired plant will be cheaper. As coal plants take more of the market, they bid up the price of coal...and so tip the balance toward hydropower. The two technologies end up sharing the market in a predictable proportion" (Arthur 1994, p.2)

Static increasing returns allow that firm which is able to capture an early advantage to remain the leader and hence the selection of a particular equilibrium to be determined by early (random) events and the system to be characterized by multiple equilibria and 'lock-in' (Arthur 1990, David 1985):

"Diminishing returns imply a single equilibrium point for the economy, but positive feedback-increasing returns-makes for many equilibrium points. There is no guarantee that the particular economic outcome selected from among the many alternatives will be the "best" one...Once random events select a particular path, the choice may become locked in regardless of the advantages of the alternatives...Predictable and shared markets are no longer guaranteed.." (Arthur 1994, p.1)

Metcalfe's (1994) exposition of the use of replicator dynamics to model internal and external economies of scale concentrates on the equilibrium and disequilibrium results from such static economies of scale.

It is argued below that static economies of scale are not adequate for understanding the dynamics of firm size and innovation for several reasons. Firstly, the term 'negative feedback' has been wrongly interpreted to produce only predictable and unique equilibria; here it will be shown that negative feedback, when used to describe dynamic diseconomies of scale, produces multiple equilibria and turbulence in market shares. Secondly, it will be argued that the dynamics of innovation are not adequately captured through the static conception of economies of scale. A quote by Kaldor (1985) refers to this second critique:

"We do not really know the causes of the uniformity in pattern which emerged in the last 50 years, under which not more than three large firms accounts for the great majority of total sales (perhaps 70-80% of the total or more) while the remainder is divided among a large number of small firms (normally several hundreds). This pattern emerged in some many different industries - like manufactures of automobiles and other durable consumer goods such as vacuum cleaners, refrigerators, electric light bulbs, or even newspapers or advertising agencies - that there must be some explanation in the dynamics of competition that goes beyond the considerations usually taken into account. Clearly, increasing returns to scale has in a broad sense something to do with it, <u>but that cannot be the whole explanation</u>, since the numbers seem to be similar in countries as different as the United States or Switzerland. One may find (I am putting this as a hypothesis) that the leading producers have the same market share in both countries, even though the size of the market is 20 to 30 times as large in one case as in the other"
(Kaldor 1985, p. 53).

In the model developed below, empirical regularities concerning market structure are reproduced and interpreted as a result not of static economies of scale, but rather of dynamic economies of scale; the effect of firm size on the rate of cost reduction. Since firms' rate of cost reduction is directly related to their innovation activities, the subject brings us to the long standing debate on whether it is small or large firms which are more innovative. Schumpeter himself had different positions on this matter, emphasizing in The Theory of Economic Development (1934) the important role of small flexible and entrepreneurial firms in the innovation process, and emphasizing in Capitalism, Socialism, and Democracy (1942) the role of large firms in financing expensive R&D activities[4]. Some economists, such as Galbraith (1956) and Kamien and Schwartz (1975), have followed more Schumpeter's second position claiming that since profit maximizing firms will only innovate if they can capture a temporary rent from innovation, it is large firms with strong market power and the necessary capital to pay for expensive R&D which will have advantages in innovation. Others, like Schumacher (1973), have instead emphasized more Schumpeter's early position.

The empirical evidence presented by Acs and Audretch (1987) and Pavitt and Wald (1971) suggests that the relation between firm size and innovation is sensitive to various factors such as the type of industry being considered, the underlying knowledge-base of the technology, and the phase in the industry life-cycle. For example, Acs and Audretch (1987) show that small firms tend to have a relative advantage in industries which are very innovative, which use a large component of skilled labor, are low in R&D intensity, and which tend to be composed of a relatively high proportion of large firms. They find that large firms instead have a relative advantage in industries which are capital-intensive, concentrated, and which produce a differentiated good[5]. Similarly to Klepper (1986) and Abernathy and Utterback (1978), they also find that small firms tend to be relatively more innovative during the early stage of the industry life-cycle when skilled labor is more important, and relatively less innovative in the mature stage when higher levels of capital intensity pose a barrier to entry for small firms. The appropriate question therefore is not:

---

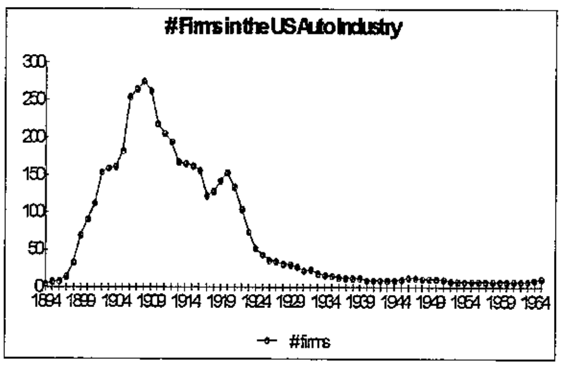
[4] *"What we have got to accept is that (the large-scale establishment) has come to be the most powerful engine of progress ..."* (Schumpeter 1942, p. 106).

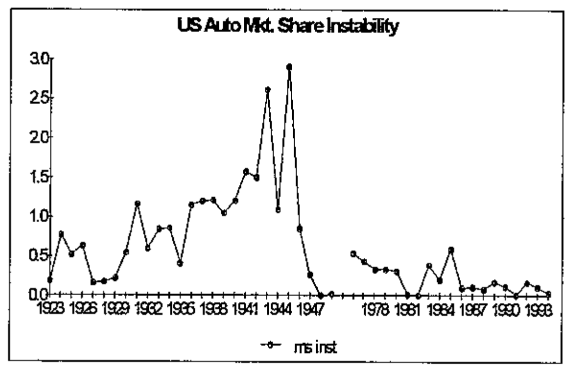
[5] They find, for example, that the innovation rate is higher for large firms in the tires, chemicals, industrial machinery and food machinery industries, while it is relatively higher for the small firms in the scales and balances, computing equipment, control instruments, and synthetic rubber industries (Acs and Audretch 1970, p. 570).

"..Which size firms have the relative advantage in innovation, but rather 'Under which circumstances do large or small firms have the relative innovative advantage?" (Acs and Audretch 1987).

Another factor, and the one most stressed in the model below, on which the relationship between firm size and innovation depend, is the specific time period in an industry's history. The industry 'life-cycle' commonly refers to the competitive situation of firms in the early, intermediate, and mature stage of an industry's evolution. The general argument is that during the early stage of an industry's history, when the product and market are uncertain, and the knowledge base is still generic, the flexibility of small new firms allows them to be the main sources of cost reduction and innovation (causing high rates of entry); while during the mature stage of the life-cycle, when the product and market demand have stabilized, economies of scale favoring large firms are strong and innovation becomes increasingly path-dependent, leading to a more stable oligopolistic structure. Hence as regards instability and concentration, in the early stage market shares are unstable while in the mature stage market shares are unstable and relatively concentrated. The two graphs in Fig. 1b below illustrating the evolution of entry and exit[6] and of market share instability (measured by the above index I) illustrate this pattern for the US auto industry:

Fig. 1b





The advantage of small size in periods of strong uncertainty, such as the beginning of the industry life-cycle, has been studied at length (Klein 1977, Jovanovic and MacDonald 1994, Klepper 1996). Uncertainty is especially high in situations of volatile changes in demand, prices and technological progress. In such situations, and in the absence of significant barriers to entry, small firms may enter the industry and due to their more novel and responsive ideas, products and processes, they challenge established firms and continuously disrupt the current way of production, organization

[6] Source for historical firm numbers: Ward's Automotive Yearbook

and distribution, wiping away the quasi-rents associated with earlier innovations. Firms with large market shares may become "lethargic", more concerned with maintaining the status quo than with initiating novel cost reduction. Hence a tradeoff between present efficiency (via static economies of scale) and flexibility might allow smaller, less specialized and more flexible firms to obtain an advantage over the larger firms. This is reversed in the mature stage of the industry when instead product standardization and specialization generate cumulative advantages for existing leaders since they generate large retained earnings (and good credit ratings) which can be used to finance increasingly expensive R&D and to hire more skilled managers and engineers. Large firms with more assets than small firms have in fact been proven to receive more favorable interest rates and loans to raise capital than smaller firms (Blitz, Bloch, Laux 1987, Siegfried 1987). Thus a cumulative pattern emerges where large firms are able to reduce their costs faster on a given set of techniques as well as to finance and search for new cost-reducing techniques (innovation). These advantages can become "barriers to entry" for small firms. Leaders remain leaders; initial advantages become permanent (absolute) advantages.

Rothwell and Dodgson (1996) summarize the advantages of large firms as 'material' advantages and those of small firms as 'behavioral' advantages. Some versions of the life-cycle argument emphasize the alternation between the above 'regimes' (rather than a uni-directional movement) as new innovations cause industrial routines to be periodically uprooted (Malerba & Orsenigo 1996).

The history of the auto industry (Abernathy & Wayne 1975, Klein 1977, Rae 1984, Kwoka 1984), the aircraft industry (Phillips 1971, Klein 1977), the television industry (Datta 1971,), the semi-conductor industry (Malerba 1985, Gruber 1994), and the tire industry (French 1990) have been related to such dynamics. Klein (1977) frames the argument in terms of periods of 'static efficiency' in which large firms have a relative advantage (exploitation of the existing production possibility frontier) versus periods of 'dynamic efficiency' in which small firms have a relative advantage (extension of the frontier). Others still have referred to the former as advantages to 'exploitation' and the latter as advantages to 'exploration' (March 1991).

This life-cycle story is used below to guide the modeling procedure. The associated relationships between firm size, innovation, and market structure, are embodied in cost equations which interact with a replicator dynamic for market shares.

IV. The Model

In what follows a simulation model is built which uses insights regarding *dynamic* (dis)economies of scale to study the origin and evolution of market concentration and market share instability. Before entering into the main exposition of the market structures which emerge from dynamic (dis)economies of scale, a brief review of static (dis)economies of scale is offered so to allow a clear comparison of the two conceptions (static and dynamic) of positive and negative feedback mechanisms.

We model the time-path of firm market shares through a replicator equation and the time-path of unit costs through a series of 'nested' equations which depict the dynamics of increasing and decreasing returns to scale. We assume an industry composed of n firms, each characterized by a market share s_i and a unit cost c_i. Although firms exit when their market shares are very low, no new firms 'enter'. Equation (1) below states that a firm's market share grows if the firm is characterized by below-average unit costs (above-average fitness) and falls if it is characterized by above-average unit costs:

$$\dot{s}_i = \gamma \cdot s_i(\bar{c} - c_i) \quad i = 1, \ldots, n \qquad (1)$$

$$\bar{c} \equiv \sum_i c_i s_i \quad = \text{average cost} \qquad (1b)$$

where $\sum_i s_i = 1$.

The parameter γ determines the speed at which firm market shares react to differences between firm efficiency characteristics. It could be made a firm-specific or an industry-specific parameter, which either remains constant or evolves over time. As an industry specific parameter, a high γ would describe an industry with a strong competitive adjustment mechanism; it might evolve endogenously with the changing Herfindahl index (as a result of selection). In the core analysis below γ is assumed constant and equal (=1) between firms, the way in which its variation affects the speed of convergence is illustrated in the case of 'simple selection'.

Equation (2) illustrates 'Fisher's principle' (Fisher 1930) which is derived from equation (1) and (1b). The theorem states that the rate of change of mean fitness is a function of the degree of variety in fitness levels across the population; the more variety exists amongst firm costs, the greater is the absolute change in the weighted average cost. Once variety disappears, the average cost remains constant.

$$d\bar{c}/dt = \sum_i c_i s_i(\bar{c} - c_i) = \bar{c}^2 - \sum_i s_i c_i^{\ 2} = \sum_i s_i \bar{c}^2 - \sum_i s_i c_i^{\ 2} =$$

$$\sum_i s_i(\bar{c}^2 - c_i^{\ 2}) = -Var(c_i) \tag{2}$$

When selection occurs amongst *constant* fitness characteristics, Fisher's principle becomes the 'fundamental theorem of natural selection' which adds a progressive element to the direction of change; the rate of *improvement* of mean behavior (rate of reduction of average cost) is proportional to the variance in population fitness. In the one-dimensional case of constant costs, the direction of change is determinate; selection leads to higher average fitness.

IVa. Static dis/economies of scale: effect of scale on the *direction* of costs

The advantage of using equation (1) to depict the evolution of market shares is that its analytical solution(s) allows us to study the dynamics of economies of scale without (for the moment) including a separate equation for costs. We illustrate this below for the case of a duopoly where s_1 refers to the market share of firm 1. The time path of firm 1's market share is derived by substituting the average cost equation (1b) into equation (1) above:

$$\dot{s}_1 = \gamma s_1(s_1c_1 + s_2c_2 - c_1)$$

$$\dot{s}_1 = \gamma s_1(1 - s_1)(c_2 - c_1). \tag{3}$$

Equation (3) allows us to study economies of scale by exploring different assumptions on the last term $(c_2 - c_1)$. We consider 3 possible cases:

(3a) $(c_2 - c_1)$	$=$ *constant*	constant returns to scale
(3b) $(c_2 - c_1)$	$= g(s_2 - s_1);\quad g'(x) > 0.$	decreasing returns to scale
(3c) $(c_2 - c_1)$	$= f(s_2 - s_1);\quad f'(x) < 0.$	increasing returns to scale

In the following we will let $g(x) = x$ and $f(x) = -x$.

We use the above three cases (3a-c) to explore the feedback mechanisms between market shares and unit costs embodied in the concept of static increasing and decreasing returns to scale.

Increasing (decreasing) returns to scale imply that as the market share of a firm increases its unit costs decrease (increase).

Case a) Simple selection

Case (3a) assumes that as the market share of firm 1 increases, the relative cost dis/advantage of firm 1 remains the same. In the evolutionary literature when agents' fitness characteristics remain constant, the system is referred to as 'simple selection'; selection occurs over a set of fitness characteristics which do not evolve with the selection mechanism (Hofbauer & Sigmund 1988). The equilibrium solution of this constant cost scenario is intuitive: the firm with the lowest cost will in the long-run capture the entire market (market share = 1). We first illustrate this equilibrium solution analytically and then via computer simulation. With constant costs, the analytical solution to equation (1) is equation (4) below:

$$ds_1 / dt = \beta s_1(1-s_1), \text{ where } \beta = \gamma * (\overline{c_2 - c_1})$$

$$\frac{ds_1}{s_1(1-s_1)} = \beta dt \Rightarrow \int_0^t \frac{ds_1}{s_1(1-s_1)} = \beta(t), \quad \text{and solving for s(t):}$$

$$s_1(t) = [1 + \frac{1-s_1(0)}{s_1(0)} e^{-\beta t}]^{-1}. \quad \text{(see Appendix A for proof)} \tag{4}$$

Equation (4) allows us to predict the equilibrium market shares of the two firms solely through our knowledge of their (relative) initial costs: as t->∞, if β is positive, i.e., if $c_2 > c_1$, then firm 1 captures the entire market, otherwise its market share goes to 0 (similarly for firm 2):

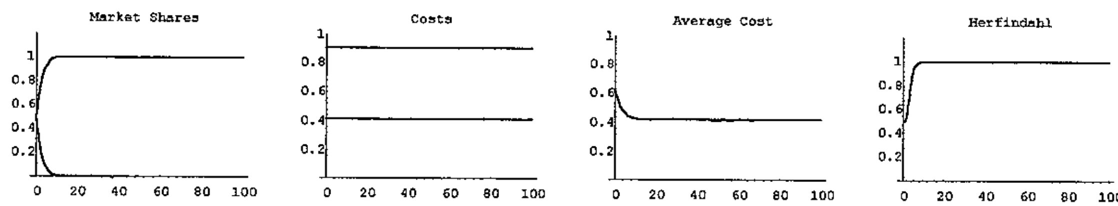
$$\text{if } \frac{c_2(0)}{c_1(0)} > 1, \text{ then } s_1(t) -> 1$$

$$\text{if } \frac{c_2(0)}{c_1(0)} < 1, \text{ then } s_1(t) -> 0$$

Thus in the case of constant costs, there is a stable long-term equilibrium of market shares: whichever firm has a cost advantage at *t=0*, will in the limit obtain a market share equal to 1. This equilibrium is independent of initial conditions; any combination of initial costs will lead to systematic divergence of the two firms' positions. The only way that both firms can co-exist is if they begin with the same costs, but then there is no evolution of market shares!

Below we illustrate the same result through computer simulation. The initial conditions are that market shares begin *equally* distributed between the n(=2) firms ($s_i = 1/n$, i=1...n) and costs begin randomly distributed (normally distributed between 0 and 1 with variance .2 and mean .6). Although costs do not change, the 'progressive' element of selection in equation (1) causes the (weighted) average industry cost to fall. We see that the firm with the lowest initial cost captures the entire market:

Fig. 2 Simulation of simple selection: equation (1) for n =2 firms with randomly distributed *constant* costs



The monopolistic result (Herfindahl index=1) holds whether firms begin with the same or with different market shares (as long as they sum to one). This is similar to the prediction in standard micro-economic theory that in the absence of diminishing returns, and/or in the presence of imperfect competition (barriers to entry, imperfect information, etc.), the leading firm will maintain its 'rent' and monopolize the industry.

The case of simple selection is not very interesting since it assumes that firms do not actively seek to reduce costs (to *change* their competitive fitness). The empirical work of 'new industrial organization' theory has countered this position by emphasizing (and documenting) the role of firm-specific actions in *altering* industry market structure through active cost-reduction strategies. The non-linear relation between market structure, firm conduct and economic performance is not compatible with the simple selection framework (Jacquemin 1985).

We now drop the assumption of constant costs and consider the case of static returns to scale, i.e. a firm's unit costs either increase or decrease with increases in market share. After deriving the equilibrium properties of static returns to scale, the core of the paper analyzes the case of dynamic returns to scale.

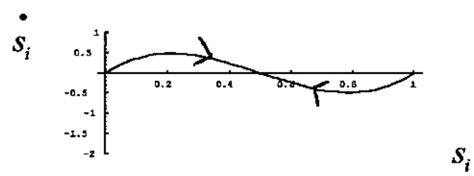
Case b) Static decreasing returns to scale (negative feedback)

Decreasing returns to scale implies that an increase in firm size has a negative effect on firm efficiency (fitness). We depict this phenomenon (case 3b) for a duopoly by substituting $(c_2 - c_1) = g(s_2 - s_1)$ into equation (3) above, where the function g ($g'(x) > 0$) illustrates that as s_1 increases, c_1 increases (increased firm size reduces firm fitness). Assuming $g(x) = x$, we may write this as: $c_2 - c_1 = s_2 - s_1 = (1 - s_1) - s_1 = 1 - 2s_1$ which transforms equation (3) into equation (5) below:

$$\dot{s}_1 = \gamma s_1(1 - s_1)(1 - 2s_1) \tag{5}$$

Figure 3a below depicts the phase diagram for equation (5). The arrows illustrate that decreasing returns to scale lead to a unique and stable market equilibrium: for *any* initial distribution of market shares, the shares will converge to a uniform value (1/n= 1/2). The equilibrium point is stable since a perturbation in either direction will cause market shares to come back to that point:

Fig. 3a



This is the standard result found in neoclassical microeconomic theory; diminishing returns to scale create negative feedback mechanisms which limit the size of a firm. It is the U-shaped average cost curve of the industry and the firm which causes the perfectly competitive market to be uniformly distributed between firms. This dynamic forms the basis for the concept of the microeconomic 'representative agent': any differential advantages between firms are competed away through negative feedback mechanisms.

As stated in the quote by Arthur above (p. 7), static decreasing returns to scale thus lead to a unique and stable equilibrium for market shares; a uniform distribution of market shares ($s_i = 1/n$, *i*=1...*n*) consistent with neoclassical 'perfect competition' theory.

Case c) Static increasing returns to scale (positive feedback)

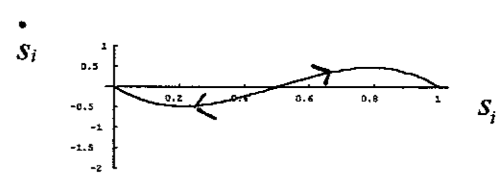
Increasing returns to scale embody positive feedback between firm size and firm efficiency: when s_i increases, c_i decreases. Learning by doing, network externalities and other phenomena leading to increasing returns allow early beginners to accumulate advantages and remain leaders. This positive feedback causes the selection of a particular equilibrium to be determined early on by small (random) events and the system to be characterized by 'multiple equilibria' and 'lock-in' (Arthur 1990, David 1985). Commonly cited examples which illustrate the selection of particular equilibrium points by early events include the competition between the Qwerty and Dworak keyboard (David 1985), the VHS and Beta standard for video-cassette recorders (Arthur 1990), and the AM and FM radio. Geographic development has also been explained through positive feedback mechanisms; Krugman (1979) claims that industrial development of the North-Eastern United States occurred due to the positive feedback for early starters who developed market networks. In international trade, forces of positive feedback have been used to explain the role of 'absolute' advantage as opposed to 'comparative' advantage (Dosi, Pavitt, Soete 1988).

To model increasing returns to scale (illustrated here for the case of a duopoly), we substitute equation (3c) [$(c_2 - c_1) = f(s_2 - s_1)$], into equation (3),where the function f ($f'(x) < 0$) states that when s_1 increases, c_1 decreases. Since $s_1 + s_2 = 1$, and by we may write (3c) as $c_2 - c_1 = s_1 - s_2 = s_1 - (1 - s_1) = 2s_1 - 1$, which when plugged into equation (3) produces:

$$\dot{s}_1 = \gamma s_1(1 - s_1)(2s_1 - 1)\,. \qquad (6)$$

Figure 3b depicts the phase diagram for equation (6). The arrows indicate that the equilibrium point (s_i =1/n) is unstable. If market shares begin equal, any small perturbation in the vicinity of that equilibrium point will cause market shares to go very far from that point: if the small movement is to the left of 1/n then $s_1 \rightarrow 0$, while if it is to the right of 1/n then $s_1 \rightarrow 1$. The system is thus characterized by multiple equilibria, where the selection of a particular equilibrium depends on initial market shares:

Fig. 3b



The dynamics of increasing returns to scale and multiple equilibria, has been widely studied in the literature on path-dependency and network externalities (Arthur 1990, David 1993, Krugman 1995) and is well represented by Arthur's quote on p. 7 above. Marshall long ago argued that if firm costs fall with an increase in market share, then that firm which gets an early advantage in share will not be limited in size: *"whatever firm gets a good start, corners the market"* (Marshall

1890). The 'good start' could be a result of luck and other non-predictable early events. Such indeterminacy causes problems of uniqueness and stability for microeconomic theory (Sraffa 1926). Brian Arthur's work has used Polya Urn processes (1985), Fokker-Planck equations (1992), and random walks (1989) to depict the dependence of the selected equilibrium on early events.

Case d) Combination of different returns to scale

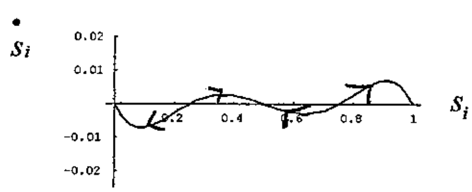
We now incorporate both increasing and decreasing returns to scale dynamics in a single cost equation, which leads to various equilibrium solutions. Equation (3d) describes a polynomial cost function which embodies increasing returns to scale within one range of market shares (between 0 and 1/*n* and between 1/*n* and 1) and decreasing returns to scale within another range of market shares (between *b* and *a*)

$$c_2 - c_1 = (1-2s_1)(s_1-a)(s_1-b), \text{ substituting (3d) into (3):} \quad (3d)$$

$$\dot{s} = \gamma s_1(1-s_1)(1-2s_1)(s_1-a)(s_1-b), \text{ with } 1/2<a<1 \text{ and } 0<b<1/2. \quad (7)$$

Figure 4, the phase diagram for equation (7), illustrates that different initial market share values will lead to different long-term equilibria. Setting a=3/4, b=1/4, and γ =1:

Fig. 4



If s_1 begins between 0 and *b*, then $s_1 \to 0$; if s_1 begins between *b* and *a* then $s_1 \to 1/n$; and if s_1 begins between *a* and 1, then $s_1 \to 1$. The points *b* and *a*, are unstable equilibrium points since very small disturbances around those points will cause strong divergence. The point 1/*n* is a stable equilibrium as long as perturbations are not very large.

IVb. Dynamic dis/economies of scale: effect of scale on the *rate* of cost reduction

Having studied the positive and negative feedback mechanisms embodied in static returns to scale, we are now ready to study these mechanisms in the case of dynamic returns to scale, which, as stated in Section III above, is considered here to provide a more sound foundation for the dynamics of firm size and innovation than the case of static returns.

To study the feedback between market structure and innovation, unit costs are assumed to always fall and the rate of cost reduction to depend on scale. Positive feedback means that an increase in market share causes the firm's rate of cost-reduction to increase (equation 8 below), while negative feedback means that an increase in market share causes the firm's rate of cost reduction to decrease (equation 9 below). In each of the simulations we assume that costs begin randomly distributed while market shares begin equally distributed ($s_i = 1/n$, i=1...n). Different variance levels of the initial distribution of costs are experimented with; this parameter can be interpreted as industry specific since, for example, it should be higher in industries where the underlying technological base is radically different from existing methods, as opposed to an industry which begins with the accumulated knowledge base from existing methods.

The simulations study how variations in the industry-specific degree of variety between firm efficiencies (initial variance of costs), and in the industry-specific speed of cost reduction (α in the cost equations below) cause different types of market structures to emerge both within and across feedback regimes. Two other factors which affect the typology of market structures are the particular draw from the cost distribution (with a given variance) and in the case of the life-cycle (cases c and d below) the time period in which a change in innovation regime occurs. We first model the case of negative feedback, then positive feedback, and then their combination in two versions of the life-cycle story. A comparison of the results to empirical literature which specifically addresses the issue of innovation and market structure using concepts very similar to the model's parameters is also included.

a) Dynamic decreasing returns to scale

We depict negative feedback between firm unit costs and market shares through a differential equation for costs (8) which illustrates that the speed of cost reduction of firm i falls as the market share of firm i increases:

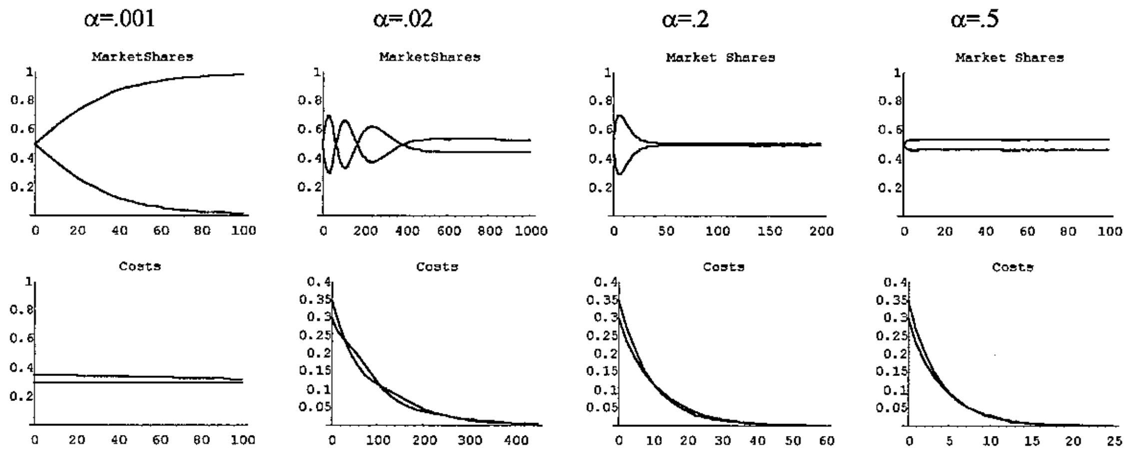
$$\dot{c}_i = -\alpha(1 - s_i)c_i \tag{8}$$

Due to the logarithmic form of the equation, the parameter α determines the speed with which intra-industry costs converge. It is an industry-specific parameter which can be interpreted as the strength of spillovers of knowledge and diffusion. The exact rate of cost reduction of any one specific firm depends both on the value of α as well as on its market share via equation (8). The value of α might depend on the 'tacitness' of the knowledge base in an industry (Pavitt 1984), on the patent system, or on other industry-specific and technology-specific factors. Due to the very different degrees of 'appropriability' and 'opportunity' conditions in industries, α should differ between industries and/or between different periods of the industry life-cycle. In the textile industry one would expect α to be higher than in the bio-technology industry since patents are less common and information is more 'codifiable'; and in the initial phase of the life-cycle of the bio-technology industry one might expect α to be higher than in the more mature phase when innovation opportunities have lessened. Although we hold α constant and look at the effect of different values of it, α might itself evolve over time as a function of the herfindahl index or of the instability index, as well as stochastically. For example, we might expect that when concentration is high (such as during the mature phase of an industry), the value of α is lower *if* concentration has a negative effect on innovation (Scherer 1984, Cohen and Levin 1989).

Before simulating the equation system describing dynamic decreasing returns to scale, intuition tells us that a turbulent 'switching' pattern of firm market shares should emerge. This is because when one firm gains a market share advantage, it's costs begin to fall at a slower rate which cause it to be passed in market share by a smaller firm whose costs are falling at a faster rate. The dynamic then repeats itself for the previously small, now growing, firm. When equation (1) and (8) are simulated below, we see in fact that a 'switching' pattern emerges and persists until the costs of the surviving firms converge to the industry average. The size of α determines the speed of convergence and hence the probability that the 'switching' has time to take place.

The time path for market shares and costs which emerge with different values of α are displayed below. The cost figures are plotted with a different time range than market shares to better display their pattern (market shares stabilize in each case a short time after costs have converged to the lowest possible cost):

Fig. 5 Simulation of equations (1) and (8) for n=2 firms with different values of α



In each case, after costs of the surviving firms converge, market shares reach their steady state value due to the dynamics of equations (1) and (1b). For different values of the parameter α, very different market structures emerge; a very high value of α causes cost convergence to occur before market share switching begins, and a very low value of α causes a monopoly to emerge before switching begins (similar to the case of simple selection). For values of α between .007 and .02, switching occurs with any initial (random) distribution of costs, but the higher the variance in initial costs, the more turbulent is the instability (measured by the Instability index). Fig. 6 illustrates this for n=10 firms[6], a given value of $\alpha = .02$ (in the switching range), and 4 different variance levels (.004, .08, .2, .4 with a constant mean of .6):

Fig. 6 Simulation of equations (1) and (8) with n=10, α=.02, and different initial cost variances

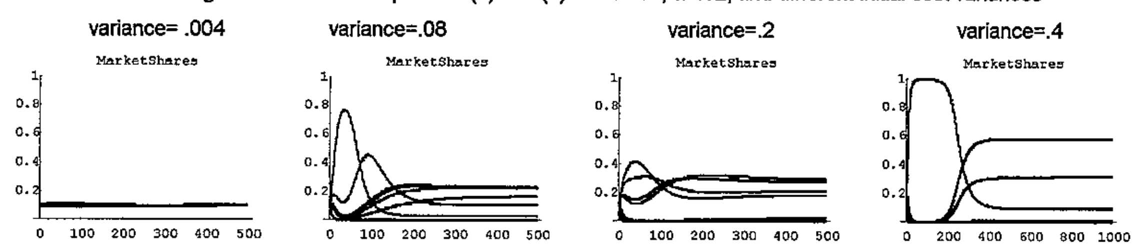
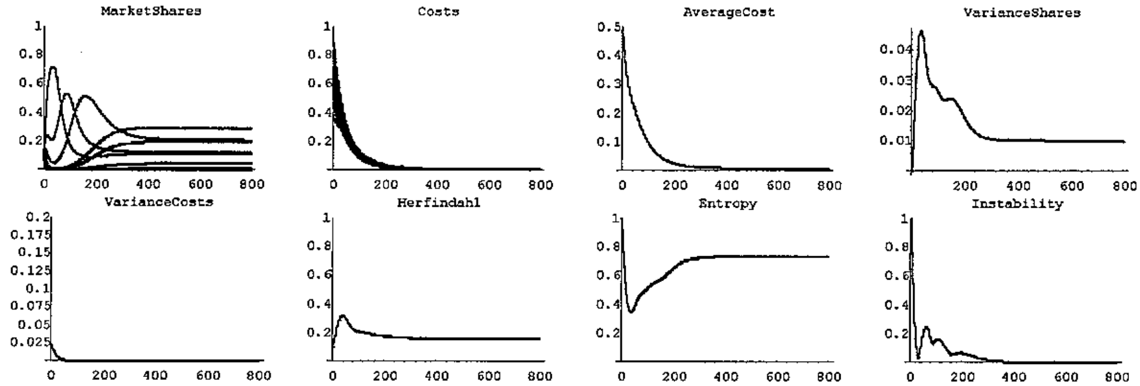


Figure 6 illustrates that a higher variance in the distribution of initial costs, causes greater market share turbulence and consequently more early exits leading to a higher degree of final

[6] When there are only 2 firms it does not make sense, due to the very small sample, to distribute costs randomly and to only specify the variance. In that case we draw from a uniform distribution with a specific range and can vary the range for experimentation. With n=10 firms we draw initial costs from a normal distribution with a specified mean and variance.

concentration. The summary statistics (including the Instability index and Herfindahl index) are found in Table 1 of Appendix B. Figure 7 illustrates the complete results for one particular case:

Fig. 7 Simulation of equations (1) and (8) with n=10, α=.02, and c_i (0) normally distributed with mean .6 and variance .2



Summary and empirical comparisons are listed below (see Table 1 in Appendix B for more detailed results):

- For values of $\alpha < .002$ a <u>monopolistic</u> market emerges. This is because when costs change very slowly (a low α implies strong inertia), selection forces completely dominate the evolutionary process allowing only the initially most fit firm (lowest cost) to survive. This is similar to the 'simple selection' case studied above. Path-dependency exists since whichever firm happens to be the leader at *t=0* will remain the leader forever.

  This result recalls empirical studies which find that concentration to be more conducive to innovation in low 'technological opportunity' industries in which the science base moves relatively slowly and predictably (Comanor 1967, Geroski 1989, Scherer 1984).

- For parameter values $.002 < \alpha < .007$, unit costs fall slowly but rapidly enough to allow partial <u>co-existence</u> of those firms (< *n*) whose costs converge before the selection mechanism has time to force all non-leader firms out of the industry. Firms with below average costs exit early. No switching or turbulence occurs since with a low α, the forces of selection embodied in equation (1) are still stronger than the negative feedback forces.

- For parameter values $.007 < \alpha < .03$, a <u>switching</u> pattern amongst market shares emerges. Negative feedback is strong enough to cause firms with increasing market shares to experience slower rates of cost reduction and to thus be surpassed in market share by smaller firms. The process causes switching to occur until all surviving firms' costs converge. For values of α closer to .03 the switching becomes so turbulent (increasing instability index) that more firms are forced to exit very early. An important result here is that, as opposed to the two extreme cases, the final ranking of firms is *unpredictable* since the initially most efficient firm does not necessarily become the industry leader.

This result recalls the empirical finding that market share turbulence tends to be higher during periods in which small firms have relative advantages in the innovative process, such as the early phase of the life-cycle (Klepper 1996).

- For large parameter values $\alpha > .03$ the switching stops since costs fall so fast that all firms reach the lowest possible cost before diseconomies of scale or selection have time to take effect. The early convergence allows all firms to co-exist and the firm with the initially lowest cost to be the final leader. The larger the value of α, the earlier market shares reach an equilibrium value. The final ranking of firms is predictably dependent on initial conditions.

  The last two results recall the empirical finding that in any given industry, market structure tends to be less concentrated in countries characterized by fast rates of innovation than in those characterized by low rates of innovation (see Dosi 1984 for the case of semiconductors). They also recall the finding that small firms have a higher innovation/employee ratio than larger firms in 4-digit industries with high innovation rates (large α) (Acs and Audretch 1987) and that vigorous innovation has been found to be more concentration reducing than increasing (Geroski 1987, Mukhopadhyay 1985). Lastly, Lunn (1986) has found that while process innovation (incremental) tends to produce a concentrated market structure, product innovation (more radical) tends to produce a more competitive market structure.

- When the simulation is run with different variance levels for the initial cost distribution (.004, .2.2 and .4, all with mean =.6), it is found that with any given value of α, the higher is the initial variance of costs, the longer it takes for market shares to reach an equilibrium value. And in the parameter range for which switching occurs ($.007<\alpha<.03$), it is found that a higher variance of initial costs, causes market shares to be more turbulent as measured by the (cumulative) instability index (see column 6 in Table 1, Appendix B), which causes more early exits and hence a higher Herfindahl index (see column 4 in Table 1, Appendix B). Thus in those industries in which the degree of variety between firm efficiencies is greater, there is a higher probability that market shares will be unstable over time, that concentration will be high, and that market shares take a longer time to reach a stable value.

  This result recalls the empirical finding that market share turbulence tends to be higher during periods in which small firms have relative advantages in the innovative process, such as in the early phase of the life-cycle or in industries characterized by strong technological vigor (Klepper 1996).

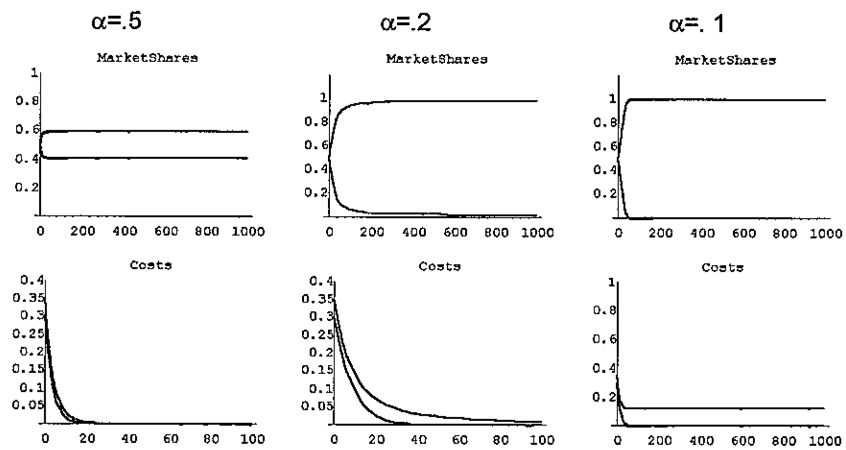
b) Dynamic increasing returns to scale

Dynamic increasing returns to scale means here that as firms grow, their capability to reduce costs increases through their greater funds for R&D, their market power to appropriate rents from innovation, etc. Equation (9) embodies this dynamic, whose economic basis was developed in Section III above, into a cost equation:

$$\dot{c}_i = -\alpha(s_i)c_i \tag{9}$$

When we simulate equation (1) and equation (9) for $n=2$ firms with initial costs randomly distributed between .3 and .4, the following time paths for market shares and costs are observed:

Fig. 8 Simulation of equations (1) and (9) with n=2



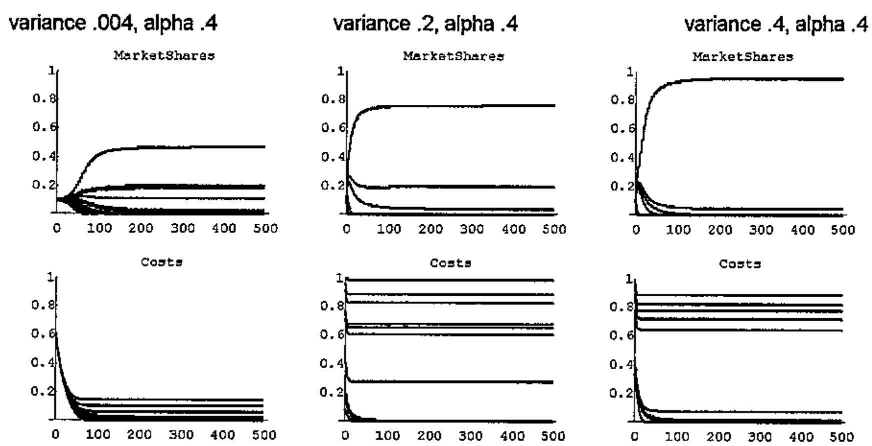
We see that a slow speed of cost reduction, i.e. strong inertia in costs (ex. due to the tacitness of the knowledge base), causes the selection mechanism to be very powerful and the 'divergence' effect to outweigh the 'catching-up' effect. In this case the asymptotic market structure has the same 'multiple equilibrium' characteristic observed in the case of *static* increasing returns observed above: firm *i*'s market share has 2 possible equilibrium points, 0 or 1, depending on initial cost conditions[1]. This result is similar to that seen in both the case of simple selection and the case of dynamic decreasing returns to scale with a low α, and hence we can say that, independently of the type of feedback mechanism (no feedback, negative, and positive), when costs fall very

[1] The same equilibrium results emerge if costs are modeled through a differential equation embodying learning curve dynamics (Mazzucato 1997). The higher is the slope of the learning curve (i.e. the faster costs decrease with cumulated output), the more firms co-exist in the market when shares reach their asymptotically stable value. The lower the slope of the learning curve, the more concentrated is the market structure which emerges.

slowly, a concentrated market emerges. As already mentioned, this seems to coincide with empirical evidence of concentrated markets in technologically laggard industries.
A fast rate of cost reduction allows a 'catching up' effect between firms to potentially outweigh the 'divergence' (path-dependency) effect characteristic of positive feedback mechanisms. Yet unlike the case of dynamic decreasing returns to scale, no value of α causes a turbulent market structure to arise; whereas in the case of negative feedback the market structure arising from a mid-level speed of cost reduction is very different from the two extreme cases, here a mid-level speed of cost reduction (α) causes the market structure to simply lie in between the two extreme cases,

When we experiment with different variance levels for the initial cost distribution we see that an increase in the initial variance of costs increases the level of market concentration, regardless of the value of α that we choose. Figure 9 illustrates this for n=10 firms with a constant value for α (= .4) and three different variance levels (.004, .2 and .4):

Fig. 9 Simulation of equations (1) and (9) with n=10, α =.4, and different variance levels (with mean .6)



The result that a higher initial variance in costs causes more early exits and hence a more concentrated market structure to emerge was also found in the negative feedback case, but there the higher variance also caused the level of turbulence to increase.

Independently of the variance level and the speed of cost reduction parameter α, the case of dynamic increasing returns to scale always makes the firm with the initially lowest cost remain the market leader. This path-dependent result where 'size begets size' (David 1984, 1985) was also found in the case of *static* increasing returns to scale and in the case of dynamic decreasing returns with with a very high or low value for α.

These results are summarized below. The empirical findings related to the simulation results are similar as those cited above in the case of dynamic decreasing returns to scale with *extreme* values of α and so are are here listed only with italicized citations.

- When the value of α is high (>.2), i.e. firm costs are falling rapidly, firms reach a stable co-existence, with their ranking determined by the initial cost advantages. The higher is α, the less concentrated is the equilibrium market structure (as α increases, the Herfindahl index approaches $1/n$). This is similar to the result for a high α in case of negative feedback.
(Acs & Audretch 1987, Geroski 1987, Lunn 1986)

- With low values of α (<.2), only one firm survives (as α decreases, the Herfindahl index approaches 1). This is also similar to the case of negative feedback.
(Comanor 1967, Geroski 1989, Scherer 1984)

- With any given level of α, the higher is the variance in the initial distribution of costs, the more concentrated is the asymptotic market structure.
(Dosi & Orsenigo 1987)

c) Life-cycle I

We now combine the above dynamics in a life-cycle scenario where negative feedback exists during the 'early' phase of the industry life-cycle, and positive feedback exists during the 'mature' phase of the industry life-cycle. The opposite is explored under 'life-cycle II' below, and can be interpreted as a third stage of the life cycle in which there is again a reversal in the fortune of the leaders. Although we do not consider 'entry' dynamics, we embody the essence of the life-cycle argument in a cost equation which claims that small firms have a faster *rate of cost reduction* than large firms during the early phase of an industry (due to their greater flexibility and adaptability to the uncertain nature of the environment), while large firms have a faster rate of cost reduction during the 'mature' phase of an industry (due to the economies of scale in R&D, and the greater role of capital intensity in production) . The more detailed economic intuition behind such a proposition is reviewed in Section III above. We incorporate such dynamics into the model by choosing an arbitrary moment in industry history $t = t_x$ at which the existing type of feedback between costs and shares undergoes a change. For our first case we assume that cost equation (8) holds from $t=0$ to $t = t_x$, and cost equation (9) holds from $t=t_x+1$ to *tmax* (the end of the simulation run). Although here we choose t_x arbitrarily it would be interesting in the future to make it a function of the industry state (the level of concentration and instability of market shares).

Given the results for the simulation of the positive and negative feedback scenarios, we expect that for values of α around .05, the simulation of the above dynamic would lead to a semi-turbulent and competitive structure during the first stage of the life-cycle and to a partially or a totally concentrated market structure during the second stage.

Because equation (8) and equation (9) embody different average rates of cost reduction, it is necessary, when setting the value for α in the two equations, not to introduce a bias into the model. The average rate of cost reduction for equation (8) is equal to $-\alpha(n-1)$, while for equation (9) it is equal to $-\alpha$. To control for this difference, the parameter α must be set differently in the two equations so that for any given value of α in equation (8), the average rate of cost reduction is the same in the two equations. We do this by replacing the average rate of cost reduction for equation (9), i.e. $-\alpha$, with the term $-\beta$. Setting the two average rates of cost reduction equal to each other we get $\beta = \alpha(n-1)$ which means that the value chosen for α in the phase of dynamic increasing returns to scale must always be $1/(n-1)$ times the value chosen for α in the phase of dynamic decreasing returns to scale. Following this rule, we set α equal to .02 in the first phase and α=.2 in the second phase. We set t_x = 200 meaning that at t=200 the industry goes from experiencing negative feedback to experiencing positive feedback between market shares and costs. Figure 10 below confirms that this dynamic leads to a period of market share turbulence followed by a period of market concentration and market share stability:

Fig. 10 Simulation with equation (8) and α=.02 from t=0-200, and equation (9) and α=.2 from t=201-1000

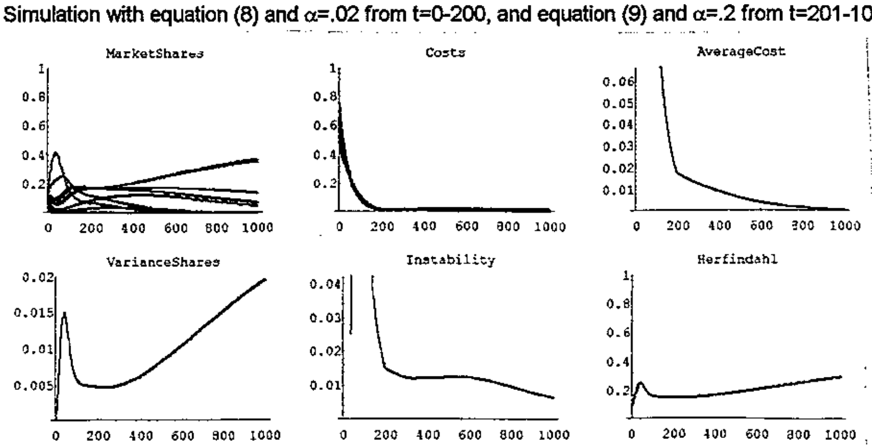
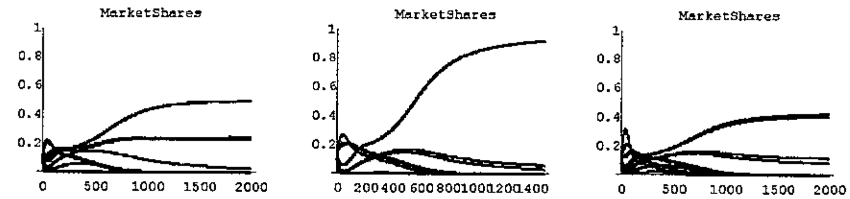


Figure 10 shows that in the period of dynamic decreasing returns to scale (from *t=0* to *t=200*) there is high and rising instability and falling concentration, measured through the herfindahl index.

Instead in the dynamic increasing returns to scale period (after *t=200*), instability falls and concentration rises.

Figure 11 illustrates that the same alternating pattern in market shares results with 3 different runs of the simulation in Fig. 10 but with different draws of initial costs:

Fig. 11 Same simulation as in Fig. 10 but with different draws on initial costs

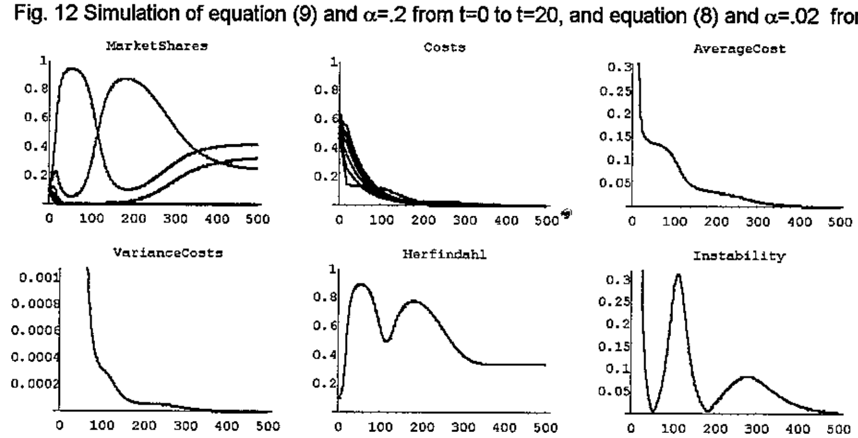


The results confirm our earlier results: as long as costs are not falling too quickly, then turbulence emerges in the case of negative feedback (seen in the first half of the left-hand side figure as opposed to the first half of the right-hand side figure).

d) Life-cycle II

We now consider the opposite case; a switch from positive to negative feedback. Here we set t_x relatively low ($t_x = 20$) since if it is set too high then total monopoly will occur very quickly and no firms will exist to experience the negative feedback phase. Again making sure not to introduce a bias into the average rate of cost reduction in the two equations, we set $\alpha = .2$ in the first phase and $\alpha = .02$ in the second phase:

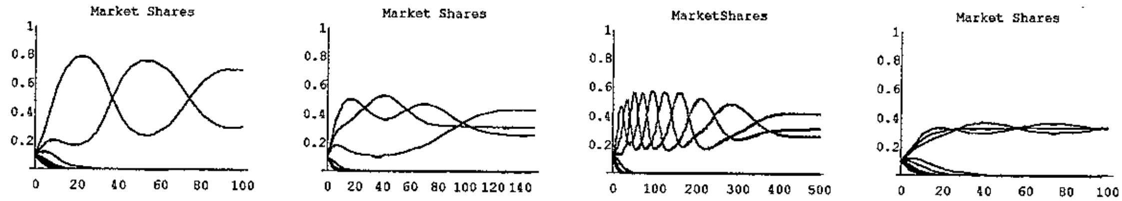
Fig. 12 Simulation of equation (9) and α=.2 from t=0 to t=20, and equation (8) and α=.02 from t=21 to t=500



We see here that in the positive feedback phase (*t=0 to t=20*) the market is concentrated, while in the negative feedback phase (t=21 to t=300) the market is less concentrated and more turbulent.

Figure 13 illustrates results from a simulation run with changes in t_x, i.e. the value for the critical time period in which the change in regime occurs:

Fig. 13 Same simulation as above but with (t_x =30, 20, 10, 5) and with α = .2 in first regime and .02 in second regime



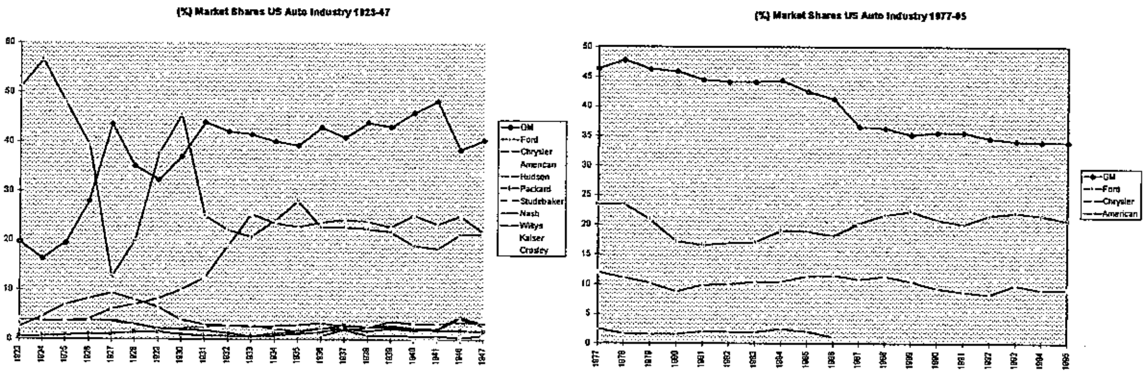
The earlier the change in feedback regime occurs, the earlier the switching occurs and the more firms are able to survive the initial shakeout.

e) Comparison of simulation results to empirical data in US auto industry

The different market share patterns linked to underlying cost dynamics which emerge from the above simulations provide some insight into market share patterns in the US automobile industry. In the US auto industry we find that dynamic increasing returns characteristic of periods in which there was little pressure for innovation (the period of mass production between 1908-1924, the 1980's) lead to more stability and concentration in market shares; while dynamic decreasing returns characteristic of periods of high innovation (the pioneering period 1890's-1908, and the immediate post-Model T period 1924-late1940's) lead to more market share instability[8]. Figure 14 illustrates these patterns (see Figure 1a for more complete view of data):

[8] We note some important phases in the history of the auto industry: (1) In the pioneering phase (1890's-1908) the uncertainty of technology, product and demand created much instability in market shares due to the importance of flexibility and exploration in discovering new techniques. This was a period of decreasing returns to specialization. (2) The advent of Ford's Model T and the system of mass production (1909-1924) allowed economies of scale to emerge where there were rewards, in terms of cost competitiveness and market share, to exploitation of existing techniques. This period of dynamic increasing returns caused the market to become more concentrated and stable. (3) When, around 1925, the market experienced an upheaval due to consumer's new taste for a closed body comfortable car, firms locked- into the system of mass production and specialization were not able to flexibly adapt to the new conditions. This created a period of market share instability where inflexible firms were penalized with large drops in market shares. Since World War 2, periods of market share concentration and stability have been characteristic of periods of low innovation.

Fig. 14 Different degrees of market share in/stability in the US auto industry

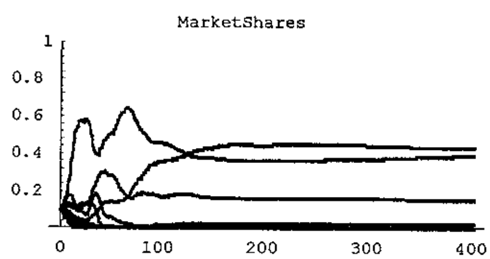


Interestingly, the type of negative feedback described by authors such as Abernathy & Wayne (1974) in their discussion of the 'limits to the learning curve' in the automobile industry, and by Klein (1977) in his discussion of static efficiency versus dynamic efficiency, is very similar to that described by our notion of dynamic decreasing returns; high increases in volume through specialization cause the *rate of cost reduction* to fall during periods of technological and market uncertainty[9]. This is also the basis of Schumpeter's statement:

> *"A system - any system, economic or other - that at every given point in time fully utilizes its possibilities to the best advantage may yet in the longer run be inferior to a system that does so at no given point of time, because the latter's failure to do so may be a condition for the level of speed of longer-run performance"*
> (Joseph Schumpeter, 1942)

Below we illustrate the similarity by comparing the above figure with market share data from the life-cycle II model. The purpose of the comparison is not to 'test' the model but to provide insight into some general processes leading to market share changes in the US auto industry. The figure below is from one particular simulation of the life-cycle II model, illustrating a first period (*t =0-20*) of positive feedback which leads to a concentrated market structure, a second period (*t= 21-100*) of negative feedback which leads to turbulent market share patterns, and a last period of market share stability when costs have converged;

Fig. 15 Simulation results from case (d) above: life-cycle I negative->positive feedback



---

[9] "*the market was changing...consumer demand began shifting to a heavier, closed body and to more comfort...the intensity of innovative activity (by Ford) diminished and the rate of cost reduction fell*" (Abernathy & Wayne 1974, p. 114).

Figure 14 and 15 allow a comparison to be drawn between the market share patterns produced by the simulation of the life-cycle I model and the actual market share patterns in the US auto industry during the selected periods in which the two types of innovation regimes are documented to exist (Abernathy & Wayne 1974). It should however be noted that instability in market shares characteristic of the earlier period does not necessarily have to come from dynamic decreasing returns as is here maintained. In fact Steven Klepper, in opposition to Abernathy and Wayne 1974 and Klein 1977, claims that no such period of decreasing returns to size ever existed in the auto industry. Rather, he claims that the instability in shares emerged from the existence of idiosyncratic events during periods of dynamic *increasing returns to scale.*

V. Conclusion and Critical Discussion

The paper presented a model in which market share patterns emerge from the feedback between firm size and costs. The model was purposefully kept simple so to: a) systematically analyze the effect of different feedback mechanisms between firm costs and market shares (scale) in isolation from other factors such as oligopolistic pricing, elasticity of demand, and entry, and b) allow the market share dynamics which emerge from the simulation to be traced to variations in few parameters which have empirical counterparts. The first reason was deemed particularly important since the emergence of path-dependency and disequilibrium dynamics from negative feedback processes is relatively unexplored (as opposed to the now numerous and very interesting studies on the generation of multiple equilibria from positive feedbacks in the economy): and the second reason was deemed important since it allows a clear difference to be posited between factors which are firm-specific, e.g. size or mutations, and those which are instead industry-specific, e.g. the rate of diffusion or initial variance of costs.

The results from this type of *Gendanken* experiment are not meant to represent the detailed dynamics of any particular industry or firm, but to produce general typologies of market share patterns linked to underlying cost dynamics. Having uncovered some *non-intuitive* properties of the deterministic dynamics, the ground is clear for the introduction of more firm-specific and industry-specific factors which will greatly enrich the dynamics. The flexibility of the model, and the fact that it does not depend on restrictive behavioral assumptions, renders it open to additions and modifications by more detailed studies. For example, the parameter denoting the speed of market share adjustment (γ) in equation (1) as well as the parameter denoting the industry specific rate of cost reduction (α) could be made to evolve endogenously with the changing

Herfindahl Index and/or Instability index. The time period in the industry life-cycle in which a change in feedback regime occurs could also be made a function of the endogenous market structure, and stochastic shocks can be introduced to represent the random external circumstances which firms to make decisions and adapt. A crucial element to any study of production and innovation is the presence of uncertainty in the development of technology and the marketing of the product, as well as the day to day uncertainty present with all production decisions. By introducing a stochastic element to costs at every period we can see how the deterministic results developed above are altered; does dynamic increasing returns to scale still result in the leadership of the initially most fit firm; how does the ability of firms to adapt to shocks *differ* in the different phases of the life-cycle?

The study connects the degree of instability and concentration in market structure to the speed at which costs are falling and the variance of the initial distribution. We review below some of the principle results:

- When firm costs fall *very slowly* (inertia), determined by a low value for the industry-specific speed of cost adjustment parameter α, the emergent market structure in *both* the cases of positive and negative feedback tends to be concentrated, with ranking depending on the initial distribution of costs (an example of path dependency). If costs are falling *very rapidly* the market structure tends to be less concentrated due to early inter-firm cost convergence. Hence, in the very slow case the force of selection overpowers the feedback mechanism causing costs to change; while in the very fast case the speed with which costs change overpowers both the effect of selection as well as the mechanism causing costs to change. The first result seems to coincide with empirical studies which have found that markets tend to be more concentrated in industries with a low rate of innovation or, in the case of a given industry, in countries which are not leaders in innovation (Comanor 1967, Dosi 1984). The second result coincides with empirical studies which find that small firms are more able to become market leaders in industries with very high rates of innovation (Geroski 1989, Scherer 1990, Acs and Audretch 1987).

- When costs fall at a medium speed positive feedback simply causes the emergent market structure to lie somewhere in between the very fast and very slow speed case, while negative feedback produces complex dynamics very different from both extreme cases. Such complexity takes the form of instability in market shares (turbulence and changes in rank) since larger firms are continuously surpassed in share by smaller firms with faster rates of cost reduction. In this case, it is not possible to predict future market shares based on current market shares. This result coincides with empirical studies which find that small firms tend to have a relative advantage in innovation activities during the early and/or more turbulent phase(s) of the industry life-cycle, when uncertainty in the environment requires flexibility and labor intensive production (Abernathy & Wayne 1974, Klein 1977).

- For each type of feedback, and with any given speed of cost adjustment parameter, the higher is the variance of the initial distribution of costs, the more concentrated is the asymptotic market structure. In the case of negative feedback with a medium speed of cost-reduction, a higher initial cost variance also causes market shares to be more turbulent. The initial variance of costs is industry-specific since it may be higher in industries where the underlying technology base is *radically* different from existing methods, as opposed to an industry which begins with the accumulated knowledge base of other methods. This result is comparable to empirical studies which have shown that low levels of asymmetry in firm characteristics cause a less concentrated market to develop (Dosi and Orsenigo 1987).

- In the 'life-cycle' case, where we embody both positive and negative feedback in the cost equation, the emergent market structure is characterized by a relatively high degree of concentration and stability in the region of positive feedback, and by a relatively high degree of instability (and varying concentration) in the negative feedback region. The exact pattern is sensitive to the speed of cost reduction parameter (α), the time period in which the 'change in regime' occurs, and the variance of initial costs. The alternating periods of instability and concentration which emerge are similar to those reported in industry life-cycle studies (Abernathy and Utterback 1975, Klein 1977, Gort and Klepper 1982, Klepper 1996).

Appendix A

Proof of equation (1b)

Equation (1b) follows directly from equation (1). For the case of n firms, we have:

(a) $\sum_{i=1}^{n} s_i = 1$ and

(b) $\dot{s}_i = \gamma s_i(\bar{c} - c_i) \quad i = 1,\ldots n$

Sum (b) over i:

(c) $\sum \dot{s}_i = \gamma \sum s_i(\bar{c} - c_i)$

From (a) we have:

(d) $\sum_i \dot{s}_i = 0$

Substitute into (c)

$\bar{c}\sum s_i = \sum s_i c_i$. Therefore, using eq. (1), $\bar{c} = \sum s_i c_i$.

Proof of equation (4) : $s_1(t) = [1 + \frac{1 - s_1(0)}{s_1(0)} e^{-\beta t}]^{-1}$.

$\dot{s}_1 = \beta s_1(1 - s_1)$, where $\beta = \gamma * (\overline{c_2 - c_1})$

$$\frac{ds_1}{s_1(1 - s_1)} = \beta dt \Rightarrow \int \frac{ds_1}{s_1(1 - s_1)} = \beta(t) =$$

$$\ln\left[\frac{s_1(t)}{s_1(0)} \cdot \frac{1 - s_1(0)}{1 - s_1(t)}\right] = \beta(t) \Rightarrow \frac{s_1(t)}{s_1(0)} \cdot \frac{1 - s_1(0)}{1 - s_1(t)} = e^{\beta t} \Rightarrow \frac{s_1(t)}{1 - s_1(t)} = \frac{s_1(0)}{1 - s_1(0)} e^{\beta t} \Rightarrow$$

$$\frac{1 - s_1(t)}{s_1(t)} = \frac{1 - s_1(0)}{s_1(0)} e^{-\beta t} \Rightarrow \frac{1}{s_1(t)} - 1 = \left(\frac{1}{s_1(0)} - 1\right) e^{-\beta t}$$

$$\frac{1}{s_1(t)} = 1 + \frac{1 - s_1(0)}{s_1(0)} e^{-\beta t} \Rightarrow s_1(t) = [1 + \frac{1 - s_1(0)}{s_1(0)} e^{-\beta t}]^{-1}$$

Appendix Bi

Table 1
Table of results for negative feedback (case (e)) in Part II with different variance levels for the initial distribution of costs (.004, .2, .4). Each value was calculated from an average of 3 simulation runs with a given value for alpha and for the variance of initial costs. The definition of each statistic is provided below the table. The values are combined in graphical form in Appendix Cii.

	1	2	3	4	5	6	7	8
	alpha	Entropy	Var. Shares	Herfindahl	Instability	Cumul. Ins.	Time Conv.	N
var = .004	0.002							
	0.02	1	0.0001	0.102	0.002	8.5	300	10
	0.05	0.99	0.0008	0.108	0.002	2.65	130	10
	0.1	1	0.0005	0.105	0.002	2.65	26	10
	0.2	1	0.0001	0.1012	0.002	1.03	17	10
	0.5	1	0.00002	0.1	0.002	0.4577	7	
	alpha	Entropy	Var. Shares	Herfindahl	Instability	Cumul. Ins.	Time Conv.	N
	0.002	0	0.089	1	0.002	21	60	1
	0.02	0.7	0.009	0.19	0.002	20	350	7
var = .2	0.05	0.75	0.011	0.21	0.002	17	180	6
	0.1	0.8	0.009	0.19	0.002	11	100	7
	0.2	0.9	0.005	0.163	0.002	6.5	35	10
	0.5	0.99	0.001	0.1	0.002	2.5	10	10
	alpha	Entropy	Var. Shares	Herfindahl	Instability	Cumul. Ins.	Time Conv.	N
	0.002	0.001	0.087	0.9	0.002	22.67	133	1
	0.02	0.5	0.1	0.35	0.002	21.84	500	4
var = .4	0.05	0.7	0.01	0.209	0.002	20.27	155	6
	0.1	0.78	0.011	0.21	0.002	9.57	42	7
	0.2	0.8	0.01	0.203	0.002	9.22	24	8
	0.5	0.99	0.001	0.116	0.002	3.44	10	10

Time of Convergence (TC) = The time at which market shares reach a steady state; determined by when the Instability Index reaches .002

N = Number of firms remaining at TC

Entropy = $-\sum_{i=1}^{n} s_i \log s_i$ calculated at TC

Relative Entropy = $E/(1/n)$ calculated at TC

Herfindahl Index = $\sum_{i=1}^{n} s_i^{\,2}$ calculated at TC

Instability Index[27] = $\sum_{i=1}^{n}\left[\left|s_{i,t} - s_{i,t-1}\right|\right]$ calculated at TC

Cumulative Instability Index = $\sum_{t=1}^{TC} I_t$ calculated at TC

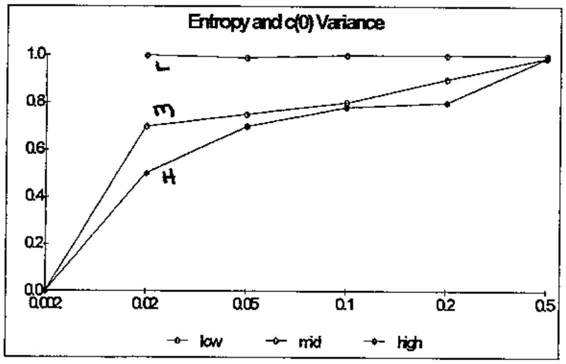
Variance of Costs = variance of costs calculated at TC

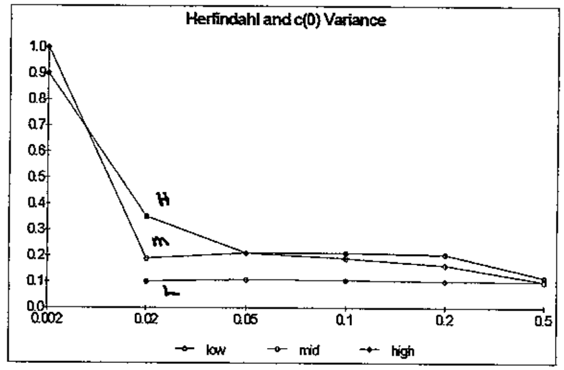
Variance of Market Shares = variance of market shares calculated at TC

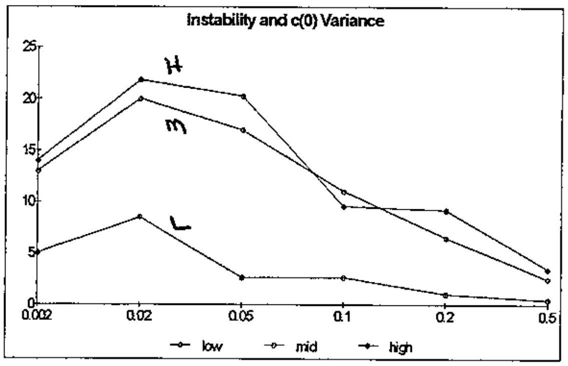
Appendix Bii

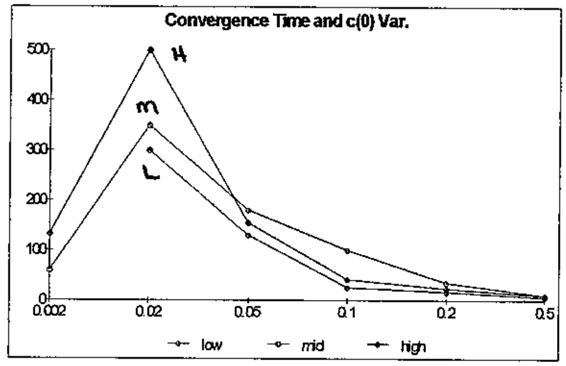
Graphical depiction of Table 1

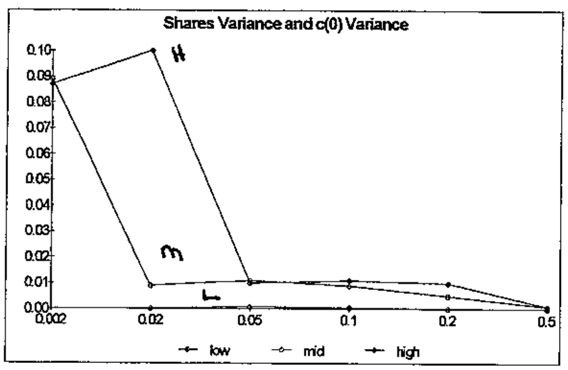
The figures below illustrate the relationship between the value of alpha and various statistics measuring competition: The three lines refer to three different variance levels of the initial cost distribution (.004, .2, .4):











References

Abernathy, W.J., Utterback J.M. (1975), "A Dynamic Model of Product and Process Innovation", *Omega*, Vol. 3, #6

Abernathy, W.J., Wayne, K. (1974) "Limits to the Learning Curve", Harvard Business Review, Vol. 52 (September 1974): 109-120

Acs, Z.J., Audretch, D.B. (1987) "Innovation, Market Structure and Firm Size", *The Review of Economics and Statistics*, Vol. LXIX, #4

Arthur, B. (1989),"Competing Technologies, Increasing Returns, and Lock-In by Historical Small Events", *Economic Journal* 99 (March 1989): 116-131

Arthur, B. (1994), Increasing Returns and Path Dependence in the Economy, Michigan Press

Audretch, D.B., (1995), Innovation and Industry Evolution, MIT Press, Cambridge, MA

Datta, Y. (1971), "Competitive Strategies and Performance of Firms in the US Television Set Industry, 1950-60", Ph.D. Dissertation, SUNY Buffalo, Business Administration dept.

David, P. (1985), "Clio and the economics of QWERTY", *American Economic Review*, Proceedings, Vol. 75: 332-7

Dosi, G. (1984), Technical Change and Industrial Transformation, Macmillan, London

Dosi, G., Orsenigo, L. (1987) "Structures, performances and change in evolutonary environments", paper presented for the IFIAS Workshop, Maastricht

Epstein, R. (1928), The Automobile Industry; Its Economic and Commercial Development, Arno Press

Federal Trade Commission, Report on the Motor Vehicle Industry, 76th Congress, First Session (1940), House Document 468

Fisher, R.A (1930), The Genetical Theory of Natural Selection, Oxford University Press, Oxford

French, M.J, (1991), The US Tire Industry , Twayne Publishers, Boston

Galbraith, J.K. (1957), American Capitalism, Hamilton, London

Geroski, P.A. ,Machin, S. (1993) "Innovation, Profitability and Growth over the Business Cycle", *Empirica*, Vol. 20:35-50

Geroski, P., Machin, S., Walters, C.,(1996) "Corporate Growth and Profitability", LBS WP

Gibrat, R. (1931), Les Inegalites Economiques, Paris, Requeil Sirey

Gort, M. (1963) "Analysis of Stability and Change in Market Shares" *Journal of Political Economy*, Vol. 62: 51-61

Gruber, H. (1994), Learning and Strategic Product Innovation: theory and evidence for the semiconductor industry, North-Holland, Amsterdam

Hofbauer J., Sigmund, K. (1988), Evolutionary Theory and Dynamic Systems, Cambridge University Press (original German edition 1984, Paul Parey Verlag Press)

Hymer, S., Pashigan, P. (1962), "Turnover of Firms as a Measure of Market Behavior", *Review of Economics and Statistics*, Vol. 44, Feb. 1962

Ijiri, Y., Simon, H. (1977), Skew Distributions and Sizes of Business Firms, Amsterdam, North Holland

Jacquemin, A. (1979), "Entropy Measure of Diversification and Corporate Growth", *The Journal of Industrial Economics*, June 1979, Vol. XXVII, #4: 359-369

Kaldor, N. (1981) "The Role of Increasing Returns, Technical Progress and Cumulative Causation in the Theory of International Trade and Economic Growth", *Economie Appliquee*, Vol. 34

Kaldor, N, (1985) Economics Without Equilibrium, University College Cardiff Press, Cardiff

Klein, Burton H. (1977), Dynamic Economics, Harvard University Press, Cambridge, Mass.

Klepper, S. (1996), "Exit, Entry, Growth, and Innovation over the Product Life-Cycle", *American Economic Review,* June 1996, Vol. 86, #3:562-583

Krugman, P. (1979), "Increasing returns, monopolistic competition and international trade", *The Journal of International Economics*, Vol. 9: 469-79

Mazzucato, M. (1997), "Three Essays on Market Share Dynamics", Ph.D. Dissertation (in progress), Dept. of Economics, New School for Social Research

Metcalfe, J.S. (1994), "Competition , Fisher's Principle and Increasing Returns in the Selection Process", in *Journal of Evolutionary Economics*, Vol. 4 (4)

Nelson, R, Winter, S. (1982), An Evolutionary Theory of Economic Change, Cambridge, Mass., The Belknap Press of Harvard University Press

Pavitt, K., (1984),"Sectoral Patterns of technical change: towards a taxonomy and a theory", *Research Policy,* Vol. 3, p.343

Phillips, A. (1971), Technology and Market Structure: a study of the aircraft industry, Heath Lexington Books, Lexington

Rae, J.B. (1965), The American Automobile, University of Chicago Press, Chicago

Rothwell, R., Dodgson, M. (1996), "Innovation and Firm Size", in The Handbook of Industrial Innovation, 1996, Edward Elgar Press, Cheltenham UK

Rothblum, U., Winter, S. (1985), "Asymptotic Behavior of Market Shares for a Stochastic Growth Model", *Journal of Economic Theory,* Vol. 36, p.352-366

Scherer, F.M., Ross, D. (1990) Industrial Market Structure and Economic Performance, Houghton

Schumacher, E.F. (1973), *Small is Beautiful,* Harper and Row, London

Schumpeter, J. (1942), Capitalism, Socialism and Democracy, Harper, New York

Schumpeter, J. (1934), The Theory of Economic Development, Harvard Univ. Press, Cambridge, MA

Schuster, Sigmund, K (1983), "Replicator Dynamics", *J. Theoretical Biology,* Vol. 100

Simon,H.A., Bonini,C.P. (1958), "The Size Distribution of Business Firms", *American Economic Review* Vol. 48(4): 607-617

US Senate (1964), Committee on Antitrust and Monopoly, Hearings, *Economic Concentration*, Part 2 (Washington: USGPO, 1964), cited in Scherer 93, p. 89

Winter, S. (1984), "Schumpetarian competition in alternative technological regimes", *Journal of Economic Behavior and Organization*, Vol. 5, p. 287-320

Young, A. (1928), "Increasing Returns and Economic Progress", *Economic Journal*, Vol. 38:527-32

View publication stats