ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Metcalfe, J. Stanley

Working Paper

On Marshallian evolutionary dynamics, entry and exit

Manchester Business School Working Paper, No. 540

Provided in Cooperation with:
Manchester Business School, The University of Manchester

Suggested Citation: Metcalfe, J. Stanley (2008) : On Marshallian evolutionary dynamics, entry and exit, Manchester Business School Working Paper, No. 540, The University of Manchester, Manchester Business School, Manchester

This Version is available at:
https://hdl.handle.net/10419/50653

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

The University of Manchester
Manchester
Business School

MANCHESTER 1824

Working Paper Series

On Marshallian Evolutionary Dynamics, Entry and Exit

J. Stanley Metcalfe

Manchester Business School Working Paper No 540

Manchester Business School

Copyright © 2008, Metcalfe. All rights reserved.
Do not quote or cite without permission from the author.

Manchester Business School
The University of Manchester
Booth Street West
Manchester M15 6PB

+44(0)161 306 1320
http://www.mbs.ac.uk/research/workingpapers/

ISSN 0954-7401

The working papers are produced by The University of Manchester - Manchester Business School and are to be circulated for discussion purposes only. Their contents should be considered to be preliminary. The papers are expected to be published in due course, in a revised form and should not be quoted without the authors' permission.

Author(s) and affiliation

Professor J. Stanley Metcalfe
Manchester Institute of Innovation Research
Manchester Business School
University of Manchester
Harold Hankins Building
Manchester
UK
M13 9PL

How to quote or cite this document

Metcalfe, J.S. (2008). On Marshallian Evolutionary Dynamics, Entry and Exit. *Manchester Business School Working Paper, Number 540* available: http://www.mbs.ac.uk/research/workingpapers/

The University of Manchester

ON MARSHALLIAN EVOLUTIONARY DYNAMICS, ENTRY AND EXIT

PROFESSOR J. STANLEY METCALFE
MANCHESTER INSTITUTE OF INNOVATION RESEARCH
THE UNIVERSITY OF MANCHESTER

This paper is a development of ideas contained in Metcalfe 2007a and 2007b. I am particularly grateful to Heinz Kurz for his very helpful comments on that early draft.

On Marshallian Evolutionary Dynamics, Entry and Exit

Introduction

Alfred Marshall was a great economist but a much misconstrued economist. Frequently portrayed as a co-founder of the neoclassical approach to economic analysis, his work is today thoroughly unfashionable, at least if leading textbooks are taken to be the standard[1]. The interests of historians of thought apart, Marshall's Principles and, *a fortiori*, his Industry and Trade repose gracefully among the apocrypha of economic thought. Who bothers today to read Marshall with all of its Victorian quaintness? I want to suggest that this attitude constitutes a misunderstanding of Marshall's contribution which was thoroughly evolutionary in tone and which bears close re-examination. Marshall's economics is a type of evolutionary industrial dynamics, in which the stationary state is a mirage, a dangerous diversion that stands in the way of a serious understanding of knowledge based capitalism; an understanding that makes the theory of value an integral part of the theory of development and economic growth. As with Schumpeter, innovation, or business experimentation as Marshall deemed it, is at the heart of this developmental scheme, and enterprise gives rise to heterogeneity in firm performance and the problem of how value is to be related to cost of production when all the firms in a given trade are different. If values depend on costs, it is not obvious whose costs we are to hold in view when we make this equation: this is a central problem for Marshall, the answer to which he gave in terms of the much maligned concept of the representative firm. In this essay I shall explore some aspects of Marshall's evolutionary dynamics, and develop a "toolkit" to help understand the relation between firm heterogeneity and the competitive process. As an application of this apparatus I will examine the role played by entry and exit in industrial dynamics, in order to demonstrate how these forces are mutually determined with the growth of incumbent firms and the growth of the market.

In this task I am emboldened by two facts, the modern revival of the "Marshall industry" (Rafaelli, 2003; Rafaelli et al, 2006), and the development of explicit, evolutionary tools of analysis that allow us to interpret Marshallian competition as a variation cum selection process that is reducible to analogous processes of replicator dynamics. Schumpeter in his semi-centennial appraisal of Marshall's work written in 1941 notably drew attention to its

[1] Schumpeter, put it thus in 1941, "In one sense Marshallian economics has passed away already. His vision of the economic process, his methods, his results are no longer ours" (Schumpeter, 1941, (1952) page 92).

evolutionary intent. There he praised Marshall handsomely as one of the first economists to realise that economics is an evolutionary science and for his contribution to formal methods of analysis. Economics certainly needs its tools but the tools are not multi purpose and so evolutionary problems require appropriately evolutionary tools. From this perspective, the central message in Marshall is that economic evolution depends on a prior notion of economic order, an order that is transient but structured. Order relates to self organisation and evolution to self transformation: to put it rather too loosely, it is why evolutionary dynamics in Marshall depends on a notion of order in the presence of diversity. Change and stasis cannot be separated. The supporting idea, which makes Marshall's approach explicitly evolutionary, is his principle of substitution in which the multidimensional heterogeneity of firms provides the material on which market framed selection processes can work their adaptive effects. Such evolutionary systems can never be at rest whilst heterogeneity is present, and entry and exit are two of the principal processes via which economic heterogeneity is created and destroyed. In this respect there is a connection between the thought of Marshall and Schumpeter: both gave prominence to the role of new entrants as vehicles to transform economic activity in individual industries.

Marshall's Approach to Entry and Exit

That the birth and death of business units is an integral part of the capitalist economic dynamic scarcely needs stating. New businesses add to capacity and, frequently but not always, they are the vehicles for innovation and the introduction of business experiments new to a trade. The disappearance of business units matters too, and, if the system is working efficiently, the firms that exit are ones whose viability has dropped below the threshold at which their resources might be better employed elsewhere. Entry and exit differ quite fundamentally though in their dynamic underpinnings. Entry is a positive decision to build capacity for a given market, it is forward looking and driven by at least an expectation of positive profits. Exit, by contrast is passive; it is a response to a loss of market and profit, it engages the management of decline when the forces at work are beyond the scope of a business to dictate its future. Moreover, we should not judge the effects of entry only in their

immediate terms, for the entrants who survive the liabilities of newness may well go on to play a significant and sometimes transformative role in an industry[2].

In Marshall, the entry and exit of businesses is one dimension of the dynamics of industry adjustment alongside the differential growth of established businesses. We first come across this theme in the Principles in relation to the famous metaphor of the trees in the forest. By this device Marshall conveys a picture of businesses differentiated by age, with the older more established businesses for ever at the risk of loosing their energy and enterprise and being overtaken by "younger and smaller rivals" (P, IV,13, p.316). But this is not a story of the independent rising and falling of different business firms. Rather the purpose of the forest metaphor is to convey the idea that the industry is an integrated ecology in which the growth of the exceptionally vigourous new entrant, the less vigorous entrants having fallen by the wayside, conditions the growth of the well established, older rivals, and vice versa. Their fortunes are jointly determined according to the relative abilities of the respective managements aided by the supporting internal and external organisation of the firms, and neither abilities nor organisation are to be taken as invariants. Marshall's explanation of entry decisions runs in terms of the anticipated profitability of establishing a particular kind of business, not its immediate profitability but rather the prospective profitability judged over a sufficiently long horizon in the expectation that the entrant will overcome any liabilities of newness and come to enjoy at least average returns for the trade. Thus expected profits are formed through an appraisal of the prevailing long period normal price and an estimate of costs in a representative, i.e. average firm in the trade in view. If the anticipated profits are greater than in other trades also open to that potential entrant he will choose that trade (P, V, 5, p.377). Of course, expectation and out turn will often transpire to be far apart, some entrants performing extraordinarily well and others failing to meet their founder's expectations, the trees in the forest again, but what matters is the prevailing incentive to add capital to a trade. The barriers to entry also vary significantly across trades and Marshall, in a rather Schumpeterian passage, makes it clear that the new man's chances of successful entry are augmented when high profits can be made through cost reducing inventions and where "by his quick resolutions and dextrous contrivances, and perhaps also a little by his natural recklessness", the entrant "forces the pace" (P, VI, 7, p.603).

[2] This theme is well documented in Baldwin (1993) and Geroski, (1995) and it is also a staple of the technology strategy and business innovation literature as, for example, in Utterback (1995) and Cristensen (1997).

It would be wrong to pretend that Marshall had tied up the loose ends in this argument, he hadn't, and it was not long before his evolutionary emphasis on business differentiation, abetted by new entrants and allied to the principle of substitution, "which is nothing more than a special and limited application of the law of survival of the fittest" (P, VI, 7, p. 597) was replaced by the Pigou/Viner equilibrium and non differentiated firm, operating in a perfectly competitive industry. Since Marshall's theory of a developmental process was replaced by a quite different account of states of equilibrium, no place could be found to explain the processes of entry and exit, they necessarily fell from view. It has been left to a small band of post-Marshallians to redevelop the process viewpoint and here P.W.S. Andrews deserves special mention. In a remarkable essay (1951), unfortunately written at a point in the development of imperfect competition theory guaranteed to ensure its minimal impact, Andrews sketched out a Marshallian theory of industry adjustment in which new entrants are given pride of place. Andrews' account is notable not least for its recognition of cross entry, that is to say, entry by firms established in other trades, as a factor adding to the entry of Marshall's new man with his own capital.

Much has been written on entry and exit in industrial economics but as Baldwin (1993) points out most of it flows from a non-Marshallian perspective on competition as a state of equilibrium. This is distinctly not the line of enquiry presented here yet there are important modern, empirical contributions which fit well with the Marshallian insights on the 'trees in the forest', industrial flux and the co evolution of a population of firms and their market. Utterback (1995) and Klepper and his co workers, in particular, have thrown a great deal of light on these phenomena under the guise of the 'shakeout' hypothesis[3]. The general observation is that entry and exit patterns follow an ordered sequence over the life of the industry as it gradually evolves towards an oligopoly. At some point in this lifecycle a drastic reduction occurs in the number of viable firms, the 'shakeout', which forms the break between a pre history and a post history of the industry. Studies of the auto, television, tyre and penicillin industries all conform to an entry, exit pattern with several common features:

- The industries in view converged to stable oligopolies over a period lasting thirty years or more, with the number of firms dropping by between 77% to 97% from their peak- the shakeout effect;

[3] Klepper (2002), Klepper and Simons (1997, 2000, 2005). The recent work by Buenstorf and Klepper (2005) puts more emphasis on fission processes in the tyre industry and finds that leading firms disproportionately generate spin-offs.

- Entry is concentrated predominantly before the shakeout but exit is distributed across the life of the industry as the forces of selection impose concentration;
- The post entry survival probabilities vary systematically with the age of firm and with the place of a firm in the entry sequence. Early entrants tend to have higher survival prospects and it is the timing of entry that most affects the expected life of a firm;
- Prior experience of an entrant in another industry, cross entry, significantly improves the probability of survival although this effect decays as the firm acquires experience of its new activities. Entrants who draw on core knowledge from other activities also have enhanced prospects of survival (Mitchell, 1989, Helfat and Leiberman, 2002)[4].

What this literature does not settle is the economic process at work, and the interrelation between the entrants, the growth of the incumbents and the growth of the market. This is what a Marshallian approach to industrial dynamics is naturally suited to do but to substantiate this claim we need to develop an evolutionary toolkit, an apparatus to capture the insights from a variation cum selection representation of economic growth. In its application we will uncover the significance of the representative firm and see how that fraught concept is an essential element in Marshall's evolutionary vision of capitalism as an open system. We shall also reflect the nature of Marshallian competition which is certainly not the perfect competition of modern thought. For Marshall, the essence of competition is the concept of economic freedom, the unfettered scope to change and experiment in business terms and, as he put it, the advantages of economic freedom "are never more strikingly manifest than when a business man endowed with genius is trying experiments, at his own risk, to see whether some new method or combination of old methods, will be more efficient than the old" (P, V, 8, 406). Two aspects of this competitive dynamic are particularly important in our exposition, the growth of incumbent businesses and the entry of entirely new businesses to a trade. Entry is often allied to innovation and it can be particularly important in the case of radical innovation but the innovative activities of incumbent firms we shall suppress in this account.

[4] Other studies confirm these general findings. Horvath *et al* (2001) apply similar methods to the US brewing industry and the shakeout that occurred in the late 19th century when the number of firms dropped by 40% in a decade. The pioneering work by organisational ecologists also provides important support to these general findings in terms of the connection of firm birth and death rates with the evolving density of a population (Singh, 1990)[4].

Evolutionary Growth as a Marshallian Process

Our apparatus is developed under two limiting assumptions that have opposing dynamic consequences. First, there are constant returns to scale at the level of the individual firm so that there are no limits to its size, only limits to its rate of growth set by the competition from rival firms and the rate of growth of the market. Implicitly, all inputs are in perfectly elastic supply at given factor prices, common to all the firms in an industry. It turns out that all the important aspects of industry evolution as an adaptive process can be explored in this particular case. Secondly, we develop the argument by the partial method in confines of a perfect product market, an important case but not the only case in Marshall's theory, in which information is so diffused among buyers and sellers as to require the output of each rival firm to be disposed of at a uniform price[5]. This has the singular consequence that the notion of a market demand curve, as distinct from the individual demand curves of particular firms, is well defined. It is an important simplification and it has definite limits from the evolutionary viewpoint, although it serves to illustrate the principle that the organisation of the market process deeply influences the dynamics of competition[6]. The more immediate implication is that it elides the difference between the customer base for a firm and the intensity with which individual customers buy the product, one of the important distinctions in Marshall's theory of demand. A perfect market also defines the fiercest of selection environments, one that facilitates the entry of new firms (by construction, customers have no loyalty to the firms they buy from), while being a harsh judge of firms on the margin of existence (there is no scope to set different prices to reflect different costs). Indeed, it is an environment that creates the maximum possible selection pressure on the population of firms defining the industry. While imperfect markets have the effect of slowing down the competitive process, increasing returns has the effect of speeding it up, so that in Marshall the former is deployed to tame the later[7]. By focusing on constant returns and a perfect market we strike a middle ground that

[5] It is perhaps worth stressing that a perfect market does not imply perfect competition, as that idea came to be formulated following Knight (1921). Marshall's evolutionary theory of competition is quite different; it is about a process not a state of affairs.

[6] In this context, it is not at all surprising that we find in Marshall's writing of the importance for market organisation and process of those innovations in transport and communication that improve the flows of information and goods, remembering that perfect markets do not mean perfect competition in the post Marshallian sense[6]. Beyond the Principles, see, for example, Marshall (1903) part II, as reprinted in Keynes (ed.) 1926.

[7] The treatment of imperfect markets is really at one with the treatment of product differentiation. For an analysis of this case see Metcalfe (1998) and for increasing returns Metcalfe (1994). Once we have imperfect markets, we enter much more into the spirit of Marshall's industry analysis in which each firm has to a degree its own particular market. Instead of a single price we have to work with a distribution of prices related to the

helps make transparent Marshall's evolutionary analysis and also keeps at bay the deeper complexities of the organic mode of evolution.

1) A Basic Marshallian Framework

We begin with the interplay between the firm and the market, and represent the industry demand curve by

$$D = N \cdot \eta(p) \tag{1}$$

N is the total number of customers in this market, each of whom purchases identical quantities of the commodity according to the intensity of demand relationship $\eta(p)$ with p being the long period, normal market price

The firm, the unit for organising production and sales, is a member of a population of otherwise heterogeneous firms in the same trade producing the same commodity. The management and organisation of the firm underpin its behavioural characteristics which are two in number, the efficiency with which production is carried out, and the investment strategy that it follows to expand or contract its capacity. Since the concern is with long period normal conditions we assume throughout that capacity and output are one and the same, their deviations being the proper subject of a short period analysis.

Market Order and the Particular Expenses Curve

Using Marshall's partial method we can arrange the given set of firms in terms of ascending levels of unit cost we can draw the long period particular expenses curve represented by the array $E - E$ in figure 1[8]. The differences between the 'steps' signify the degree of unit cost variation in the population and the lengths of the 'steps' signify the inter-firm differences in

distribution of unit costs and other firm characteristics, albeit on the same lines as explored below. The more general case is, of course, more interesting but little of substance is lost in this prologue by starting from the special conditions of "perfect" markets. It is with respect to product innovation that this claim begins to lose force, which is one good reason why, Marshall fashion, I have locked such novelties into the pound of ceteris paribus.

[8] By the term 'long period' we mean only to require that all costs, prime and supplementary, are included in the construction of the particular expenses curve. We do not mean by this phrase that all possible adjustments to profitable investment opportunities have been achieved: the long period method applies to a process not to a state of equilibrium.

installed capacity[9]. The curve is a product of the industry's history, of the pattern of accumulation, of entry and exit and of many of the creative forces that we have provisionally locked in the pound ceteris paribus, particularly innovations in organisation and technique. Of course its 'position and shape' are transient and greatly influenced by the rates of entry and exit as we shall establish below.

Given a perfect market, a demand curve such as $D-D$ may be imposed on the diagram and the resultant point of co ordination defines the normal, long period market order characterised by a price and a "marginal firm" or group of firms with the same unit costs. Marginality is described by unit cost level m with the demand curve. $D-D$, and any firm with this cost level exactly breaks even taking account of prime and supplementary costs. However, with demand curve $D'-D'$ no firm is marginal although firms with unit cost level f are the closest to a state of marginality. Any firm with unit costs higher than the ruling normal price is deemed to be non viable, to have ceased production and to have scrapped its capacity- hence the dotted form of the curve beyond $D'-D'$. Thus there are two possible configurations of a market order to contend with, and the differences between them play an important role in the analysis of industry evolution. Case I corresponds to curve $D'-D'$, in which the long period market order is maintained through price adjustment. It is the flex price case. Case II corresponds to the demand curve $D-D$ in which long period market order is maintained through quantity adjustment, that is to say, by changes in the utilization of the capacity of marginal firms selling at a price equal to their long period average costs. This is the fixed price case in which any reductions of capacity are concentrated in the population of marginal firms. The long period order captured in figure 1 is the story Marshall tells in book V. Marginal activities, if they exist, just break even and all the other active firms earn surplus profits or composite, long period rents on their past joint investment in technique, organisation and knowledge. It is the surplus profits earned by the infra marginal firms that play the central role in the following account of Marshallian long period competition for they are the basis for the growth of some incumbent firms and the stimulus to entry.

[9] It would be straightforward to interpret each step in terms of a number of equally efficient establishments owned by different firms, and to allow any one firm to be located at different points on the particular expenses curve. The equating of firms and establishments in the following is a convenient but innocuous simplification. We also hold constant throughout the factor prices that underpin the particular expenses curve. The focus in this paper on the changing length of the steps reflects the decision to suppress innovation and the forces changing the height of the steps.

Incumbent firm dynamics are captured in a simple rule, a classical rule that links investment in new capacity to profitability but which allows for idiosyncratic variations in investment strategy. This rule is as follows. If the operating return is greater than a particular threshold value the firm expands at a rate determined by the finance it can raise. If the rate of return is less than the threshold value the firm does not invest, it is stationary although it may still be profitable. If the rate of return drops below zero, the firm exits the industry.

To formalise this, let g_i be the firm's growth rate of capacity and output, and let the accumulation rule be expressed as [10]

$$\begin{aligned} g_i &= -\phi_i + \mu[p - h_i] \quad ; \quad (p - h_i) > \phi_i / \mu \\ &\leq 0 \quad ; \quad \textit{otherwise} \end{aligned} \tag{2}$$

In this expression, h_i denotes unit cost in the firm and the ratio ϕ_i / μ defines for each firm its critical investment margin, equivalent to Marshall's "outer limit, or margin, of profitableness" (P, V, 4, p.356). The coefficient ϕ_i is interpreted as a strategic investment threshold, a measure of a firm's willingness to invest at a given rate of return, while the coefficient μ, common to all firms, we call the investment coefficient. The smaller is ϕ_i the more expansive is the investment strategy. We assume that each firm distributes profits to the owners and draws investment funds from the 'capital market' at a rate that depends on its profitability. Finance constrains growth, and the fact that μ is the same for all firms means that the capital market is not discriminating between firms with the same operating profit[11]. We should notice immediately the asymmetrical nature of this investment strategy rule. The growing firm obeys (2) but the stationary or declining firm obeys a quite different set of rules in relation to the utilisation of existing capacity and the decision to scrap capacity that is underutilised. This is a crucial Marshallian distinction, the economics of decline are not the

[10] In Marshall, the notion of capacity depends on much more than the plant and facilities in place, it depends on organisation and on what modern theorists would call the distributed capabilities of the employees. Investment is to be interpreted in these broad terms.

[11] We have also assumed, for simplicity of exposition, that the capital-output ratio is the same for each firm. In modern terms we might say that the firm's cost structures differ in a Harrod neutral way.

economics of expansion in reverse and this asymmetry plays an important role in the following analysis.[12]

In figure 1 we can incorporate this new dimension of firm behaviour if we add the dotted, stepped line, labelled h^*, lying above the particular expenses curve to show the normal price for each firm that corresponds to its critical investment margin. In the case of cost level f and demand curve $D-D$ we see any such firm is profitable but that the prevailing long period price is below the critical value required to justify investment. Such firms are profitable but stationary. It seems intuitively clear from figure 1 that the relation between profitability and expansion in the aggregate depends on the joint distribution of the two characteristics ϕ_w and h_w, and so it will turn out to be. The significance of allowing the firms to follow different investment strategies is that two firms with the same unit costs and profit margin may grow at very different rates in terms of capacity and market shares. While real world competition is certainly many dimensional, a framework with just two of many possible dimensions of inter firm variety is sufficient to provide the richness we need to explore and develop Marshall's theory of competition in evolutionary terms.

The immediate consequence of combining the organisation of firm and market in this way is to highlight the possibility that a particular, active firm can occupy one of several different, mutually exclusive categories according to its characteristics and the prevailing environment and, moreover, that it can change the category it is located in over time. There are three possibilities[13]:

- Dynamic firms that operate within their investment margin and are growing in absolute terms but may be further subdivided according to whether they are rising or falling in relative size. In addition to these two dynamic states we have

[12] It could be claimed that this is the essence of the distinction between long period and short period forces, one aimed at expansion, the other at decline or even stationarity. Asymmetry of investment behaviour is important to our argument Negishi (1989, p. 367), for example, derives the result that price equals average cost in the representative firm only by assuming that all firms follow an identical investment rule and that the process of decline in capacity is identical to the process of expansion apart from the direction of the change. Symmetry leads directly to the idea that price equals cost in the representative firm but we are not told how the necessary presence of loss making firms is compatible with long period conditions.

[13] Of course, other possibilities can extend the list, for example, the firm that sets prices less than unit costs and covers the deficit either by drawing down liquid reserves or benefits from subsidies from a beneficent banker or state. The general point that this hides is the importance in modern capitalism of the factors that shape the non viability of firms such as bankruptcy and insolvency law.

- Stationary firms that are profitable but to a degree that is less than the amount required to induce them to invest, and so neither grow nor decline absolutely but do decline relatively to the rest of the (growing) market; and, finally,
- Marginal firms that have ceased to earn positive profits and are normally contracting absolutely and relatively and heading for exit from the industry; it is from this group of firms that we deduce the exit rate measured in terms of capacity reduction.

Of the three groups it is the dynamic group that drives the evolutionary process of Marshallian adaptation and it is this group that is the focus of the long period method[14]. The stationary group may account for a large share of the total market but they do not impose development on the industry. For the marginal firms too, investment strategy is of no relevance, rather their problem will be how to manage the relation between reductions in demand and reductions in capacity to define the exit rate for the industry. As we shall show below, the rate of exit is determined by the difference between the natural growth of the market, the rate of entry, and the growth of incumbent firms. Indeed, new entrants if they are viable at entry must also fall within one of the three classes defined by the marginal, stationary and dynamic groups. The group within which any one firm may fall is not an intrinsic property of that firm but a derived consequence of the firm's engagement with the competitive market process. The intrinsic characteristics of a firm, its efficiency and investment strategy, matter greatly but only relative to the characteristics of rival firms and the nature of the market environment.

To identify how the population of firms is distributed across the three groups, we introduce the device of a selection set and the operation of partitioning that set. The selection set is simply another way of representing the data from figure 1 in relation to the particular expenses curve and the threshold prices that determine investment behaviour. Partitioning is the way that we introduce Marshall's flux into the depiction of evolution.

A Marshallian Toolkit: The Selection Set and Partitioning.

[14] Were we to allow for increasing returns, then the dynamic firm would also be the host to any internal economies that are denied to the stationary or marginal firm. See Appendix H in Principles for the implications of irreversibility in the evolution of costs.

The selection set is a primitive concept in this evolutionary model. It is defined by the whole population of firms that are subjected to the same causal, market forces in relation to demand, capacity and their growth. In principle, it may have any number of dimensions but in our terms it is defined by the two variational characteristics, $\phi_{\dot{w}}$ and $h_{\dot{w}}$. We represent this in figure 2 by the convex space defined by the set of points $\{s\}$ and the linear boundaries that connect them[15]. It is entry of firms that has created this set and any firm that has existed in this population since the foundation of the industry is represented by a point in this space, the boundary and interior of which may be sparsely or densely populated. In fact only two firms are required to define a meaningful selection set. As we explore below, the nature of the set will change as new firms enter, possibly redefining the boundary, and as existing, active firms innovate by changing unit costs or investment strategy. For the moment let us hold all such changes in abeyance and explain the idea of an evolutionary partitioning of a given selection set.

Partitioning

It is the prevailing long period normal price taken from figure 1 that provides the first step in the partitioning, for it separates the viable from the non viable firms. By erecting the locus AV at p on the unit cost axis in figure 2 we deem all the firms located to the right of this locus to be inactive. Cases I and II differ only in respect of the fact that, in the latter, some active, marginal firms are positioned on the AV locus. To the left of AV will lie the dynamic and stationary firms. It is the former, combined with any new entrants, that will drive the adaptive process, for it is only this group that invest and in so doing reshape the particular expenses curve and, depending on the co-movements of the demand curve, redefine the long period order. Any consequential changes in the long period, normal price will change the nature of the partitioning and transfer firms between the different possible categories. It is in this way that the evolutionary dynamics is premised on the prevailing market order, so figures 1 and 2 are interconnected and our next task is to exploit this fact and further refine the partitioning and separate the dynamic from the stationary firms.

To do so we need the following measures of the population structure. Let s_i be the share of a dynamic firm i in the total output of the dynamic group of firms, with s_j and s_k defined as

[15] Any point on this boundary represents a firm which is a market share weighted linear combination of the firms defining the end points.

the corresponding shares for a stationary firm and a marginal firm in their respective groups. Let e be the share of the non-dynamic firms in total output, and f be the share of marginal firms in the output of the non–dynamic firms[16]. It follows that we can define the growth rates for each class, as $g_s = \sum s_i g_i$; $g_c = \sum s_j g_j$; and, $g_m = \sum s_k g_k$. Since the growth rate of the stationary firms is by definition equal to zero, we can write the aggregate growth rate of the industry's output as[17]

$$g = (1-e)g_s + efg_m \tag{3}$$

The growth rate g_m is the proportional rate of reduction of capacity by marginal firms, while the magnitude efg_m measures that reduction in capacity as a fraction of the total output of the industry. It is this later ratio that we call below the exit rate.

In normal conditions the rate of growth of industry capacity must be equal to the rate of growth of total market demand which is given, from (1), by,

$$g_D = g_N - \alpha g_p \tag{4}$$

Where g_p is the rate of growth of the long period normal price, g_N is the natural rate of growth of the number of customers in this market, and α is the elasticity of market demand. Relations (3) and (4) capture the fundamental features of the evolution of the industry, sans entry, in which either the rate of change of the normal price or the rate of change of capacity in marginal firms maintains order over time.

The distinction between the two cases of market order discussed above now becomes important and we begin by developing the flex price case, case I, in which there is neither

[16] The shares in total market demand and in the different sub populations are defined by $d_i = (1-e)s_i$; $d_j = e(1-f)s_j$, and $d_k = efs_k$

[17] In case I, the growth rate of the marginal firms is zero and (3) should be adjusted accordingly.

exit of marginal capacity nor entry of new capacity by producers new to the industry. The analysis, therefore, is conducted entirely in terms of the dynamic and stationary firms.

To identify the boundary between the stationary and dynamic firms we need only equate the firm growth rate in (2) to zero to define the locus $A-A$ in figure 2, with slope of -μ, which cuts the horizontal axis at a value equal to the prevailing long period normal price. By construction, any firm located on or above this locus is stationary although it will be profitable. All the dynamic firms are located below this locus. In this way, a knowledge of the current long period, normal price provides the basic partitioning of the industry into non viable firms in area N, marginal firms on the locus AV, stationary firms in the area C and on the locus $A-A$, and dynamic firms in the areas, D. Once we have identified the sub-population of dynamic firms we can use the output shares s_i to define the average values of unit costs and the propensity to accumulate as $\bar{h}_s = \sum s_i h_i$ and $\bar{\phi}_s = \sum s_i \phi_i$ respectively.

Because the stationary and marginal firms have no active role in the accumulation process, we can focus on the dynamic firms, and if we aggregate equations (2) across this group, using the weights, s_i, we find that their average growth rate is given by

$$g_s = -\bar{\phi}_s + \mu\left[p - \bar{h}_s\right] \tag{5}$$

By using (5) to eliminate p from (2), the growth rate of each dynamic firm, in normal conditions, can be expressed in terms of the deviations of ϕ_i and h_i from their corresponding dynamic group population averages, thus

$$g_i = g_s + \left(\bar{\phi}_s - \phi_i\right) + \mu\left[\bar{h}_s - h_i\right] \tag{6}$$

Equation (6) is fundamental to the evolutionary interpretation of Marshall's theory, since it captures the different contribution that each dynamic firm makes to the reshaping of the particular expenses curve and to changing the structure of the industry. It also embodies a distance from mean dynamic or replicator process that is the signature of a variation cum selection based evolutionary processes. The expansion of any one firm relative to the growth of capacity as a whole depends on how its characteristics compare to the population averages:

a higher than average efficiency and a greater than average propensity to invest support growth at a rate greater than the group average.[18] The relative rising and falling of firms also captures Marshall's principle of substitution in which the growth rate of each firm is mutually determined with the growth rates of its rivals. The growth rate diversity that ensues is the clue to Marshall's evolutionary dynamics.

Equating (6) to zero we can derive an alternative expression for the locus $A-A$ in terms of the characteristics of the dynamic firms and their deviations from average behaviour in the dynamic group. Taking account of (3) we can write the stationary-dynamic boundary as

$$\phi_i = \frac{g_D}{1-e} + \bar{\phi}_s + \mu\left[\bar{h}_s - h_i\right] \qquad (7)$$

There is one further element in the partitioning to be uncovered, that which separates rising from falling firms, and this is where the representative firm makes its first appearance.

The Representative Firm

Where in the selection set can we locate the representative firm? To answer this question we have to deal with two problems. The first problem relates to the fact that our firms are differentiated in two dimensions, creating the possibility that representative behaviour serves only to define a trade-off between unit costs and propensities to accumulate. The second problem is deeper, and connects to the very idea of Marshall's industries as populations of firms. By a population, I simply mean a set of entities constituting an ensemble. By an evolutionary population, I mean a set of entities whose changing relative importance in the set, however measured, is the caused outcome of specific selection and other causal processes operating equivalently on all the members of that population. Marshall's industries in this view are evolutionary populations unified by the action of common market forces. Hence the problem: a representative firm is only representative relative to some evolutionary population, and if we change the population, or sub population, in terms of its constituent firms or the causal processes acting on them then we change necessarily what is meant by representative.

[18] Notice though that the relation between efficiency and growth depends on the investment coefficient and thus on the capital: output ratio and our treatment off the capital market.

While there is much ambiguity in Marshall's discussion of a representative firm, we draw on two of his characterisations to define a representative firm as some average firm that maintains constant its share in the output of the relevant population. From this we see immediately that two notions of representativeness come to the fore. One is Marshall's representative firm that grows as quickly as the total market, a definition that makes no distinction between the different categories of active firm, whether dynamic, stationary or marginal. Marshall's representative firm maintains constant its share in *total market capacity and demand for the industry*. The other is a dynamically representative firm, one that grows as quickly as the population of dynamic firms in the population[19]. The dynamically representative firm maintains constant its share in the *total output of the dynamic group.* The two notions of what it is to be a representative firm are quite different and only equate to one another when the only active firms are in the dynamic category. One can see immediately, that, in general, the dynamically representative firm must expand more rapidly than Marshall's representative firm, which, of course, is growing at the same rate as the industry as a whole.

To locate the dynamically representative firm we proceed as follows. We have already identified the relevant population as the sub-set of firms located below the locus $A-A$ in figure 2. However, this ensemble of dynamic firms consists of two sub groups, occupying the regions labelled D_1 and D_2, according to whether the firms in question are rising or falling in terms of their relative outputs. To separate these two groups is to identify the dynamically representative firm. Because such a firm is growing at the average rate for the dynamic sub-population, we set $g_i = g_D/(1-e)$ in (6), to define a locus labelled $R-R$ in figure 1b. This locus passes through the point of population means for the investment and unit cost characteristics of the dynamic sub population. The equation of this locus is independent of the market growth rate and the structure of the active population, and it is given by

$$\phi_i = \bar{\phi}_s + \mu\left[\bar{h}_s - h_i\right] \tag{8}$$

[19] By the same token there is a representative stationary firm but we leave this aside since stationary firms play a minor role in the dynamics of the industry.

It also has a slope of $-\mu$, and it stands vertically below $A-A$ by a distance that measures g_s, the growth rate of the dynamic group. Like $A-A$, the locus $R-R$ is a statistical construction, a dividing line, in this case, between two classes of dynamic firms. Any actual firm whose unit costs and strategic investment stance place it on this locus will be growing at the same rate as the dynamic group as a whole, and so may be labelled dynamically representative. Either side of the $R-R$ locus we find Marshall's flux. Any firm in the region D_1 between this locus and $A-A$ is expanding absolutely but declining relatively to the output of the dynamic group as a whole. Any firm below $R-R$ in the region D_2 is expanding absolutely and relatively to the dynamic group[20]. It follows from this definition that dynamically representative behaviour is described in terms of the capacity share weighted average characteristics of the firms in the dynamic group. The definition of these averages is not arbitrary but flows from the underlying theory. However, *pace* Marshall, 'representative-ness' does not define a single firm but rather a whole possible family of them, in fact those hypothetical firms that lie on the locus $R-R$.[21] Of course, no actual firm need lie on this locus at any one time, nor does our account require that there is such a firm. This locus simply defines the possibility that any firms that should happen to be on it will be growing absolutely but neither rising nor falling relatively in the dynamic group. This is the dynamic significance of the representative firm, a significance that is entirely lost in a static treatment of competition as a state rather than as a process.

If $A-A$ identifies the stationary-dynamic boundary, we might now enquire which of the firms in the selection set will have the highest long period growth rate. Since growth rates increase as we move toward the origin from $A-A$, it follows immediately that it is firm b on the South-West boundary of the selection set that is the fastest growing, and is thus the dynamic firm on which the output of the industry is concentrating the fastest. Given the shape of the selection set, the location of this firm depends only on the parameter, μ, that is to say, on the assumed characteristics of the capital market and the investment requirements for expansion of the firms. It is immediately apparent that, in general, the fastest growing

[20] While $R-R$ passes through the point of dynamic population means, it must be remembered that these means depend on the partitioning between the dynamic and other groups, and thus vary in general as the normal long period price varies.

[21] Moreover, there is a further property hidden in this statistical concept of representativeness. For the point defined by the population means is also a point that corresponds to the average profit margin in the dynamic group. To the left and above this point any representative firm has a higher than average profitability and to the right and below it has lower than average profitability.

firm is not the least cost firm in the selection set, for this is firm a in figure 2. As in any variation cum selection analysis, the competitive dynamic selects for the characteristics that maximise expansion (the analogue of evolutionary fitness) not the characteristics that only maximise efficiency. Indeed, only if all firms followed the same investment strategy would efficiency be maximised in this adaptive process[22].

We have already indicated that our dynamically representative firm is not, in general, Marshall's representative firm. Where now do we locate Marshall's representative firm, that firm that maintains its share in the total market rather than any one sub market? Necessarily it too must be a dynamic firm, for it must be growing if the market is growing, though it is necessarily growing less rapidly than the dynamic representative firm we have just identified. To identify its properties we set $g_i = g_D$ in (6) to define the locus labelled $M-M$ in figure 2. This locus is given by the equation

$$\phi_i = \frac{g_D \cdot e}{(1-e)} + \bar{\phi}_s + \mu\left[\bar{h}_s - h_i\right] \tag{9}$$

This locus lies in region D_1 above the locus $R-R$ and below the locus $A-A$, its position depending on the share of non-dynamic firms in the overall population of active firms, on the investment coefficient, and on the overall growth rate of market demand. Any firm on this locus is maintaining a constant share in total industry capacity and thus market demand but has a falling share of the capacity within the dynamic group of firms. Where this locus cuts the unit cost axis is defined by the point $\tilde{h}^*$ in figure 2[23]. It should now be clear why it is only in an industry consisting entirely of dynamic firms that the two concepts of the representative firm coincide. As soon as we allow for non-dynamic firms they diverge, and it is the notion of the dynamic representative firm that we must give primary attention too, for it is the dynamic firms that are the carriers of long period adaptive forces. The dynamic

[22] In which case the selection set would collapse to a horizontal line the end points defining the most efficient and least efficient firms. This is what Marshall's discussion is usually interpreted to imply. It will be noted that the direction of maximum growth is always orthogonal to the $A-A$ locus

[23] The vertical distance between $A-A$ and $M-M$ measures the market growth rate, g_D. A crucial difference compared with the dynamically representative locus $R-R$, is that the whole industry population means for ϕ_i and h_i do not lie on $M-M$ except by fluke. The vertical distance between $M-M$ and $R-R$ measures the value of eg_s. The share of the dynamic group of firms in total output $(1-e)$, is measured by the ratio of the distances AM to AR.

representative firm, therefore, acts as the fulcrum around which the industry is evolving and around which all of the flux in the industry can be rendered intelligible.

Thus the idea of a representative firm, in either form, is derived from the idea of a long period partitioning of the selection set. Partitioning provides a complete characterisation of the rising and falling of firms, their viability and non viability. Indeed, figures 1a and 1b are complementary ways of representing the relation between a market order and its self transformation. Partitioning is also the way we give content to the idea of economic structure not only in terms of relative scale but also in terms of dynamic characteristics of different firms and groups of firms. This is Marshall's point, the industry is always co-ordinated but the firms are all different, in terms of growth or decline and whether they are dynamic, stationary or marginal. It is from this heterogeneity that the progress of the industry is derived and the degree of heterogeneity will be shaped by processes of entry and its effects on exit.

2) Entry and Exit Processes

We can now put this toolkit to work by investigating the impact of entry and exit on the evolution of the industry. We do so under the assumption that the lifetime characteristics of an entrant are fixed at the time of entry, there are no post entry changes in efficiency or strategy. Marshall did not make such an assumption but to allow for changes after entry would take us into the territory of innovation and change of investment strategy by incumbent firms which is beyond our current brief.

To begin we need to distinguish between the entry and exit of firms, the units of economic control, and the entry and exit of productive capacity, the units of activity. Occasionally the two coincide, as in Baldwin's (1993) notions of Greenfield entry and closedown exit, but more generally they do not. A firm, for instance, may exit the industry by selling its assets as a "going concern" to another incumbent firm, to leave capacity unchanged with one fewer firm, or it may sell to a firm outside the industry so that the entrant of a new firm does not result in a change of capacity. The reason for dwelling on this distinction is that the immediate evolution of the long period market order is determined by changes in capacity relative to changes in market demand and so it is in capacity terms that we must formulate the impact of the entry and exit of firms. However, the longer term evolutionary impacts depend

on the characteristics of the firms that enter and the firms that exit the industry. That entry and exit are not mirror images of one another relates to the asymmetry, hinted at previously, that the loss of capacity through exit is a passive process, while the addition to capacity caused by the entry of a new firm is an active investment decision. It is a distinction between reacting to events and creating events. Once a firm is reduced to the marginal category, its fate is determined by the rate of expansion of the dynamic group relative to the natural rate of growth of the market. By contrast, entry is a dynamic determinant of the evolution of the industry and two dimensions of the entry problem matter, the location of the entrant in the selection set in terms of its unit costs and investment strategy, and the scale of entry in terms of the immediate addition to industry capacity.

The fundamental point to grasp is that entry creates the selection set; it is in this sense the fundamental determinant of the scope for economic evolution in the industry. Yet the immediate effect of entry may mask its more significant consequences, namely to transform the selection set and the future course of evolution. We may term a radical entrant one that redefines the boundary of the selection set (figure 2) and its SW border in particular, possibly creating a new best practice firm to displace firm b as the focal point for the concentration of the industry. This may be the kind of entry to associate with fundamental technical innovation and an associated investment strategy that outperforms the existing incumbents. We can contrast it with the possibilities for normal entry, in which the entrant is positioned within the prevailing selection set and its subsequent fortunes depend on its location relative to the prevailing partitioning of that set. On the one hand, it may be a loss making or marginal entrant located on or to the right side of AV with costs equal to or greater than the prevailing long period normal price, in which case its life will be short. On the other hand, if the entrant is profitable it may be positioned either among the stationary or the dynamic groups of firms. There is nothing unusual in a stationary, profitable entrant; this is quite a normal case, although it means that its relative importance in a growing industry is forever in decline. However, it is the dynamic entrants that matter in terms of the future transformation of the industry for they are the entrants that expand capacity after entry. Here we can distinguish between entrants in regions D1 and D2 who respectively decline and expand relatively, and, clearly, it is the entrants in region D2 which have the greatest sustained effect on the industry. We may also note that, comparing two firms that enter at the same date, it is the firm that has the lowest unit costs that will have the longer economic life but not necessarily the greater economic impact on the industry.

By treating the consequences of entry in terms of the selection set and its partitioning we can relate post entry performance to investment strategy as well as to productive efficiency. In this way we can pinpoint the effects of different entry rates on the long period normal price and on the wider pattern of industry evolution

Entry, Exit and Capacity Accounting

The first step is to account for the effects of entry and exit on total capacity. Consider a short interval of time Δt, and let n be the increment of capacity generated by new entrants in that interval when expressed as a proportion of the industry output at the end of the interval. Let d be defined analogously as the proportion of industry capacity that is eliminated through exit in this interval. Notice that the exit of capacity, to repeat, is not equivalent to the exit of firms. Define g_c as the growth rate of the continuing, incumbent firms (stationary and dynamic) and, as above, let g be the overall growth rate of capacity output. It is an elementary matter of accounting that these ratios are related by

$$(1+g)(1-n+d)=1+g_c$$

Or, when expressed in continuous time by

$$g = g_c + n - d$$

If there are marginal firms, it follows that the exit rate is given by $-d = efg_m$ and that $g_c = (1-e)g_s$, so that the market order preserving condition becomes

$$g_D = g_N - \alpha g_p = (1-e)g_s + efg_m + n \qquad (10)$$

Two possibilities follow. In case I there are no marginal firms and so exit in capacity terms is zero, and $f = -d = 0$. In case II, some firms are in a state of marginality, they are reducing capacity, and f >0. We may note in passing that the available empirical evidence suggests that entry and exit rates are typically small in absolute terms, apart from the early stages in a

new industry they are usually less than 5% of capacity (Baldwin, 1993, 2003). However, this is not the point that matters, it is the scale of entry and exit relative to the growth of the market that determines the immediate evolution of the industry and these two numbers may be of comparable magnitude.

What of the entry rate? From a broad perspective we should acknowledge that the opportunities for entry may depend on wider possibilities for innovation, and treat the entry rate as a parameter, noting its possible volatility in terms of the number of entrants and their initial scale of operation. But it has generally been agreed that the entry rate bears a positive relation to the prevailing profitability of the industry, albeit as modified by the manner in which current profitability is translated into expectations of future profitability. If so, entry and the performance of incumbent firms will be jointly determined and in a fashion that depends on assumptions about the link between entry and profitability. This is certainly as Marshall saw it and so, while recognising that the entry decision may reasonably depend on many factors, it is the relation of entry to current profitability that is the focus of economic attention. In an industry of differentiated incumbents this poses a problem, namely, "Whose profitability is to be taken as the guide to entry?" Since the incumbents differ in profitability, from the high of best practice to the low of marginality, Marshall's answer was expressed in terms of the profitability of the representative firm, which we interpret as the average profitability for the group of dynamic firms[24]. Moreover, just as the incumbents differ in their characteristics so do the entrants, and higher levels of incumbent profitability may be expected to induce the entry of less adventurous or less efficient newcomers[25]. We can formalize this Marshallian argument as follows.

Each potential entrant (w) judges the case for entry by comparing its prospective profitability at the ruling price against a threshold unit profit margin denoted by φ_w. Entry occurs if expected profitability, exceeds this entry threshold, and expected profitability is taken to be the profitability of the dynamic group of firms, $\left(p - \bar{h}_s\right) = m_s$. Among the distribution of

[24] This corresponds to the profitability of the average dynamic firm, that firm with average characteristics on the locus $R - R$ in figure 2. However, on this locus, the profitability of any representative firm increases as we move up the locus from right to left towards firms with lower unit costs.

[25] The barriers facing entrants are not independent of their treatment by the capital market. Our understanding of entry for new small businesses in particular reflects the possibility that they are not treated on equivalent terms to incumbents when it comes to raising capital. For a useful survey of venture capital aspects of this phenomenon see Freeman (2005)

potential entrants the most likely entrant defines a threshold profit margin below which entry will not occur, φ_0 , and, for margins greater than this, the entry rate is correspondingly greater but less than proportionally so. This we can express in an entry function[26]

$$n = \sigma\left(p - \bar{h}_s, \varphi_0\right) \tag{11}$$

the general form of which is shown in figure 3. Special cases include one in which entry is independent of the profitability of the representative firm, the entry function is horizontal at the going entry rate, as well as the Andrews case (1951), in which the list of potential entrants is so large and homogeneous that the entry function is vertical at the threshold, φ_0. In this later case, the long period price is determined by the entry process alone and it varies *pari passu* with unit costs in the dynamically representative firm.

In figure 3 we can develop a characterisation of the immediate impact of entry on profit margins and thus long period values in terms of our cases I and II. The figure is drawn holding constant the characteristics of the representative firm. Consider first case II as depicted in figure 3a. Here the normal long period price is "locked" by costs in the prevailing marginal firms, to produce the average margin m_1 for the dynamic incumbent firms. The growth of the market equals the natural growth rate g_N as shown by the horizontal line. From this line we deduct the contribution to industry growth made by the dynamic sub population of firms, the schedule labelled g_s with slope $(1-e)\mu$. At the prevailing value m_1, the growth contribution of the incumbents is measured by the distance bz and the rate of entry by the distance $m_1 n$. Since incumbent growth and entry add up to a rate of capacity expansion that exceeds the natural market growth rate, we have a positive rate of capacity exit measured by the distance nz.

This exit rate is given by

$$-d = g_N - \sigma\left(h_m - \bar{h}_s, \varphi_0\right) - \mu(1-e)\left(h_m - h^*\right) \tag{12}$$

[26] We could just as well relate the entry rate to the margin in Marshall's representative firm discussed above, adding the average unit cost in the stationary group to the determinants of the entry rate. This would not add anything to the dynamic discussion other than to provide an additional factor to shift the entry function in figure 3.

Where, $h^* = \bar{\phi}_s / \mu + \bar{h}_s$, is defined by the given characteristics of the representative firm, while h_m measures the ruling long period market price. As can readily be established, any variation in the entry rate corresponds to an equal and offsetting variation in the exit rate so that the change in net entry is zero; as Geroski (1995) pointed out, it is the case where entry induces exit[27].

Case II prevails for as long as marginal firms are active but once they are eliminated we enter the world of case I in which the long period price adjusts to balance the growth of demand with the growth of capacity. This case is depicted in figure 3b which again takes the characteristics of the dynamically representative firm as given. We may first ask, "Is there a configuration of the growth rates consistent with a constant long period price, a position in which $g_D = g_N$?" There is and it is denoted by point z, where, given the value of h^*, the jointly determined contributions of entry and of incumbent growth add up to the natural growth of the market. It is intuitively clear that the effect of entry is to sustain a lower margin and thus a lower growth rate in the dynamic firms than would be the case without entry. What happens when the system is not in this quasi steady state, for example, when it is characterised by the non steady state margin m_2? This is a situation in which the price is declining and so the market growth rate line g_D, lies above the line g_N by the magnitude $-\alpha g_p$[28]. In this situation the growth contribution of the incumbents is measured from the market growth rate line along the schedule with slope $(1-e)\mu$. At the prevailing margin m_2, the entry rate is $m_2 n'$, the incumbent growth contribution is $b'z'$ and their sum exceeds g_N the natural rate of market expansion. Consequently, the distance $n'z'$ measures the compensating rate of expansion of the market due to the required rate of decline in the long period normal price. It will be clear that price reduction and exit play symmetric roles in this process of industry evolution, and that entry, exit and incumbent growth rates are mutually determining.

[27] Striking empirical confirmation of this matching entry and exit phenomena is contained in Peter Murmann's authorititve account (2003) of the development of the synthetic dye industries in Germany the UK and the USA, in which growing cumulative entry rates are matched by growing cumulative exit rates. As Murmann puts it, "More failures and more successes seem to have been parallel processes". *Op cit*, p.43.

[28] It is a straightforward matter to determine the rate at which the market price is changing but we do not pursue this question here. For details see Metcalfe 2007b.

The account we have given of the entry process is Marshallian particularly in regard to the role of the average profit margin that is earned in the dynamic firms. However, it is a partial approach, and it is clear that this margin will change over time as a result of entry and the expansion of incumbent firms, including past entrants. To give a complete account of these interlinked processes would take us beyond the scope of this paper. Moreover, these calculations also leave aside the many other factors which transform average cost levels, innovations, induced transfers between the dynamic and stationary groups that flow with repartitioning the selection set, and the relative rise and decline of firms due to competitive selection. But evolution is a complicated process and Marshall was right to claim that we gain some understanding by taking one problem at a time.

Conclusion

Marshall's Principles is framed around a vision of capitalism as a dual process of self organisation and self transformation, it is an open system in which the theory of value becomes an integral part of the theory of economic development. The processes of self transformation are closely linked to innovation and business experimentation, and the entry of new firms in a trade is one of Marshall's principal vehicles for inducing market adaptation to the possibilities latent in the generation of economic variety. In this Marshall has much in common with Schumpeter, another great economist of vision and indeed a remarkably similar vision of the role of novelty and profits in economic progress. The interesting difference is that Marshall developed tools to analyse economic order and its transformation. In this essay I have sort to spell out a Marshallian evolutionary toolkit to focus on entry and exit, two of the many kinds of economic flux that are found in the evolution of industries. They are important signatures of economic evolution in their own right and entry has the added role of acting as a vehicle for innovation and even more fundamental economic transformation. Marshall famously chose the phrase "natura non facit saltum" as the epigram of the Principles and in so doing lured generations of economists into a belief that his vision had no dynamic substance, that all the talk of evolutionary change and biology was misguided prattle. Not so. I hope the reader will agree that there is more to Marshall than meets the eye.

Bibliography

Andrews, P.W.S, 1951, 'Industrial Analysis in Economics – With Especial Reference to Marshallian Doctrine', in T. Wilson and P.W.S. Andrews (eds), Oxford Studies in the Price Mechanism, Oxford, Clarendon Press.

Baldwin, J.R, 1993, The Dynamics of Industrial Competition: A North American Perspective, Cambridge, Cambridge University Press.

Baldwin, J.R., 2003, Innovation Strategies and Performance in Small Firms, Cheltenham, Edward Elgar.

Buenstorf, G. and Klepper, S., 2005, 'Heritage and Agglomeration: The Akron Tire Cluster Revisited', mimeo, Evolutionary Economics Group, MPI, Jena.

Cristensen, C.M., 1997, The Innovator's Dilemma: When Technologies Cause Great Firms to Fail, Harvard University Press.

Freeman, J., 2005, 'Venture Capital and Modern Capitalism', in V. Nee and R.Swedberg (eds), The Economic Sociology of Capitalism, New Jersey, Princeton University Press.

Geroski, P.A., 1995, 'What Do We Know About Entry', International Journal of Industrial Organisation, Vol. 13, pp. 421-440.

Helfat, C.E. and Leiberman, M.B., 2002, 'The Birth of Capabilities: Market Entry and the Importance of Pre History', Industrial and Corporate Change, Vol. 11, pp. 725-760.

Horvath, M., Schivardi, F. and Woywode, M., 2001, 'On Industry Life Cycles: Entry and Shakeout in Beer Brewing', International Journal of Industrial Organisation, Vol. 19, pp. 1023-1042.

Klepper, S., 2002, 'Firm Survival and the Evolution of Oligopoly', RAND Journal of Economics, Vol. 33, pp. 37-61.

Klepper, S. and Simons, K.L., 1997, 'Technological Extinctions of Industrial Firms: an Enquiry into their Nature and Causes', Industrial and Corporate Change, Vol. 6. pp. 379-460.

Klepper, S. and Simons, K.L., 2000, 'The Making of an Oligopoly: Firm Survival and Technological Change in the Evolution of the US Tire Industry', Journal of Political Economy, Vol. 108, pp. 728-760.

Klepper, S. and Simons, K.L., 2005, 'Industry Shakeouts and Technological Change', International Journal of Industrial Organisation, Vol. 23, pp. 23-43.

Knight, F., 1921, Risk, Uncertainty and Profit, Boston, Houghton Miflin Co.

Marshall, A., 1903, 'Memorandum on the Fiscal Policy of International Trade, Part II', reprinted in J.M., Keynes (ed) (1926), Official Papers of Alfred Marshall, Macmillan, London.

Marshall, A., 1919, Industry and Trade, Macmillan, London.

Marshall, A., 1920, Principles of Economics (8th edition), London, Macmillan.

Metcalfe, J.S. 1994, 'Competition, Fisher's Principle and Increasing Returns to Selection', Journal of Evolutionary Economics, Vol. 4, pp. 327-46.

Metcalfe, J.S., 1998, Evolutionary Economics and Creative Destruction, London, Routledge.

Metcalfe, J.S., 2007a, 'Alfred Marshall and the General Theory of Evolutionary Economics', History of Economic Ideas, Vol. 15, pp.81-110.

Metcalfe, J.S., 2007b. 'Alfred Marshall's Mecca: Reconciling the Theories of Value and Development', The Economic Record, Vol. 83 (Supplement), pp. 1-22.

Mitchell, W., 1989, 'Whither and When? Probability and Timing of Incumbent's Entry into Emerging Industrial Subfields', Administrative Science Quarterly, Vol. 34, pp. 208-230.

Murmann, J.P., 2003, Knowledge and Competitive Advantage, Cambridge, Cambridge University Press.

Negishi, T., 1989, History of Economic Theory, Amsterdam, North Holland.

Raffaelli, T., 2003, Marshall's Evolutionary Economics, London, Routledge.

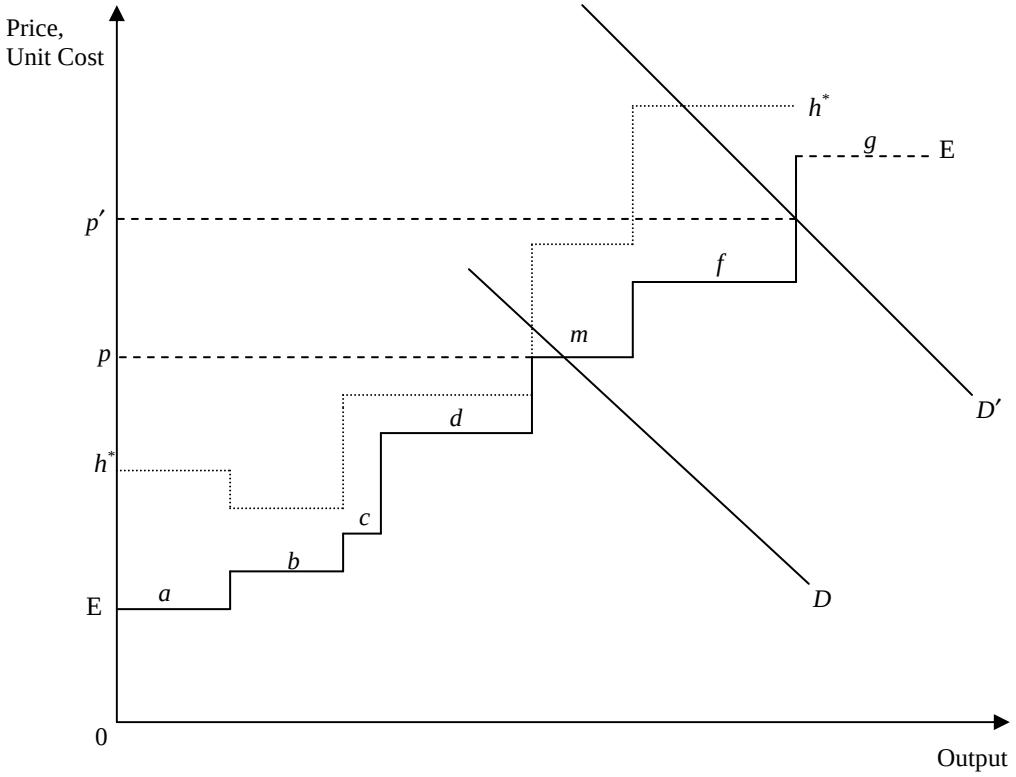
Raffaelli, T., Becattini, G. and Dardi, M. (eds), 2006, The Elgar Companion to Alfred Marshall, Cheltenham, Edward Elgar.

Schumpeter, J.A., 1941, 'Alfred Marshall (1842-1924)' in J.A. Schumpeter, 1952, Ten Great Economists, London, George Allen & Unwin.

Singh, J.V., 1990, Organisational Evolution: New Directions, London, Sage.

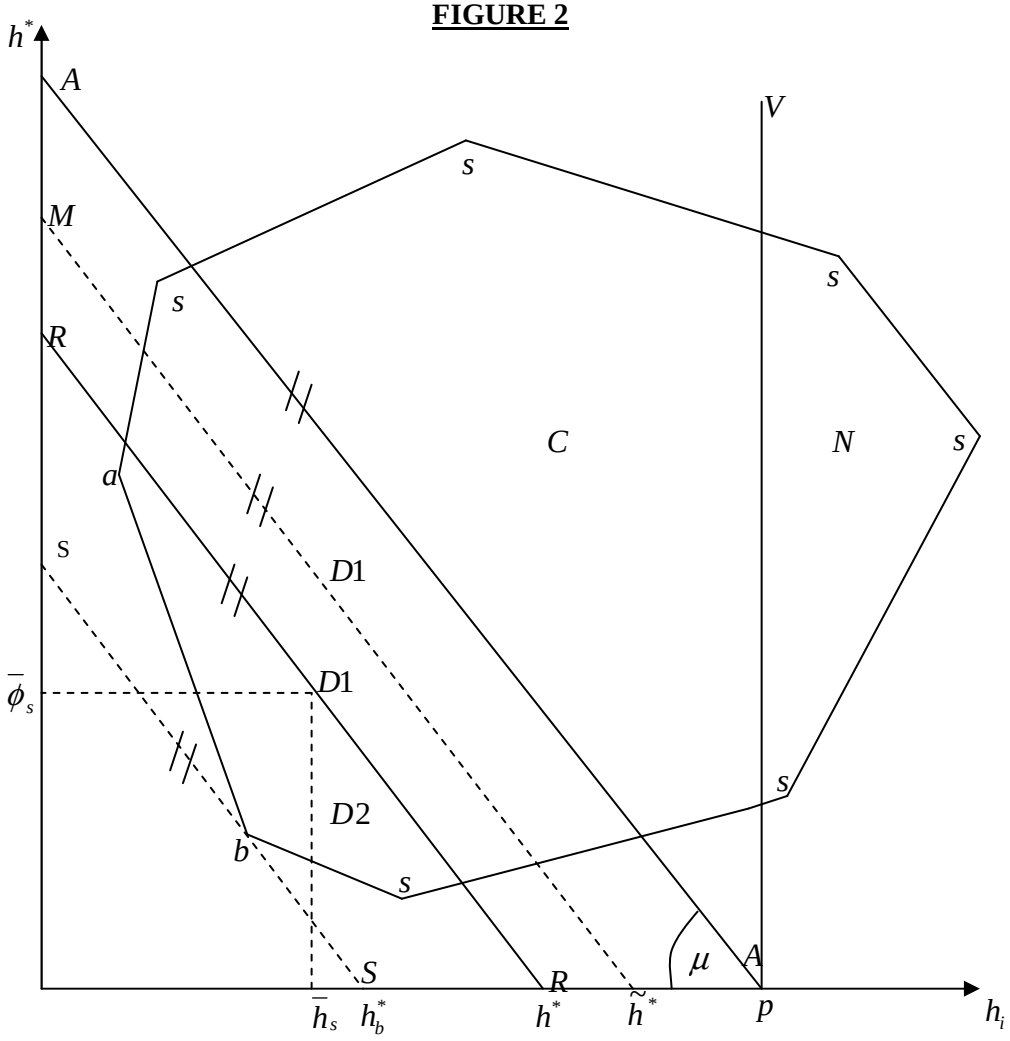
Utterback, J., 1995, Mastering the Dynamics of Innovation, Harvard University Press.

FIGURE 1



The Augmented Particular Expenses Curve

FIGURE 2



The Selection Set

Figure 3a

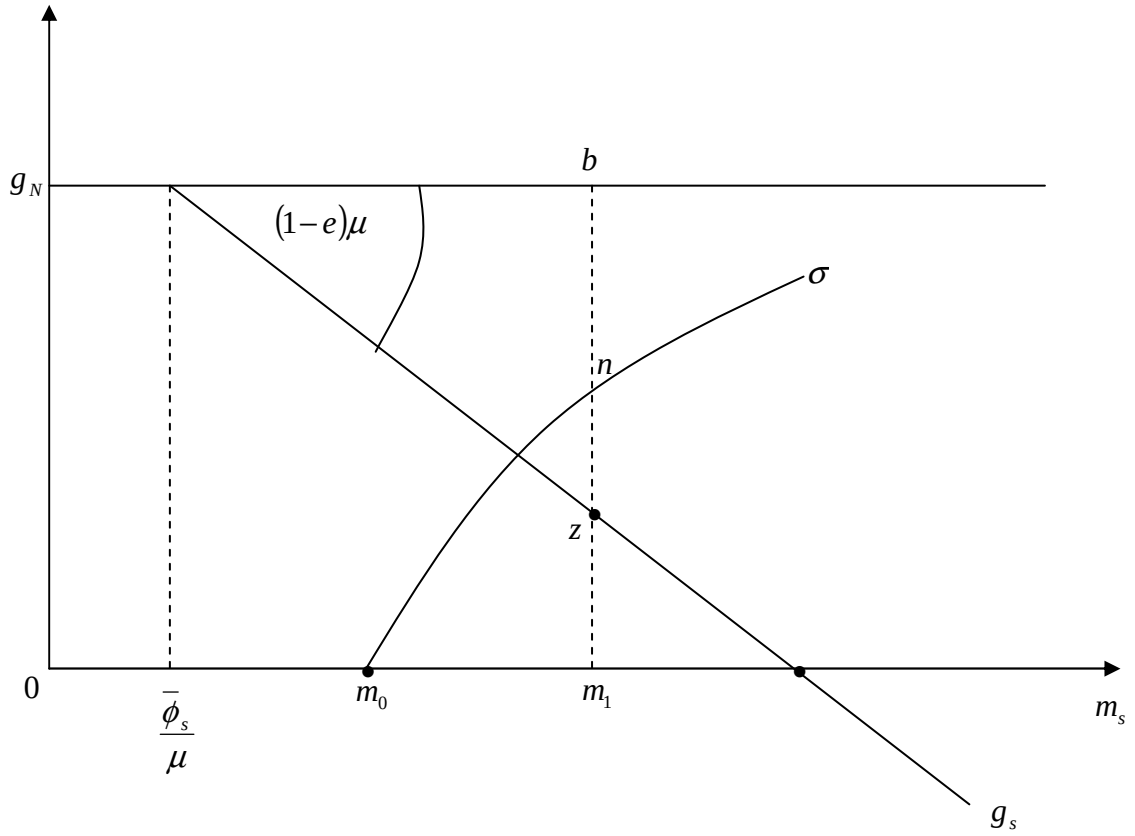
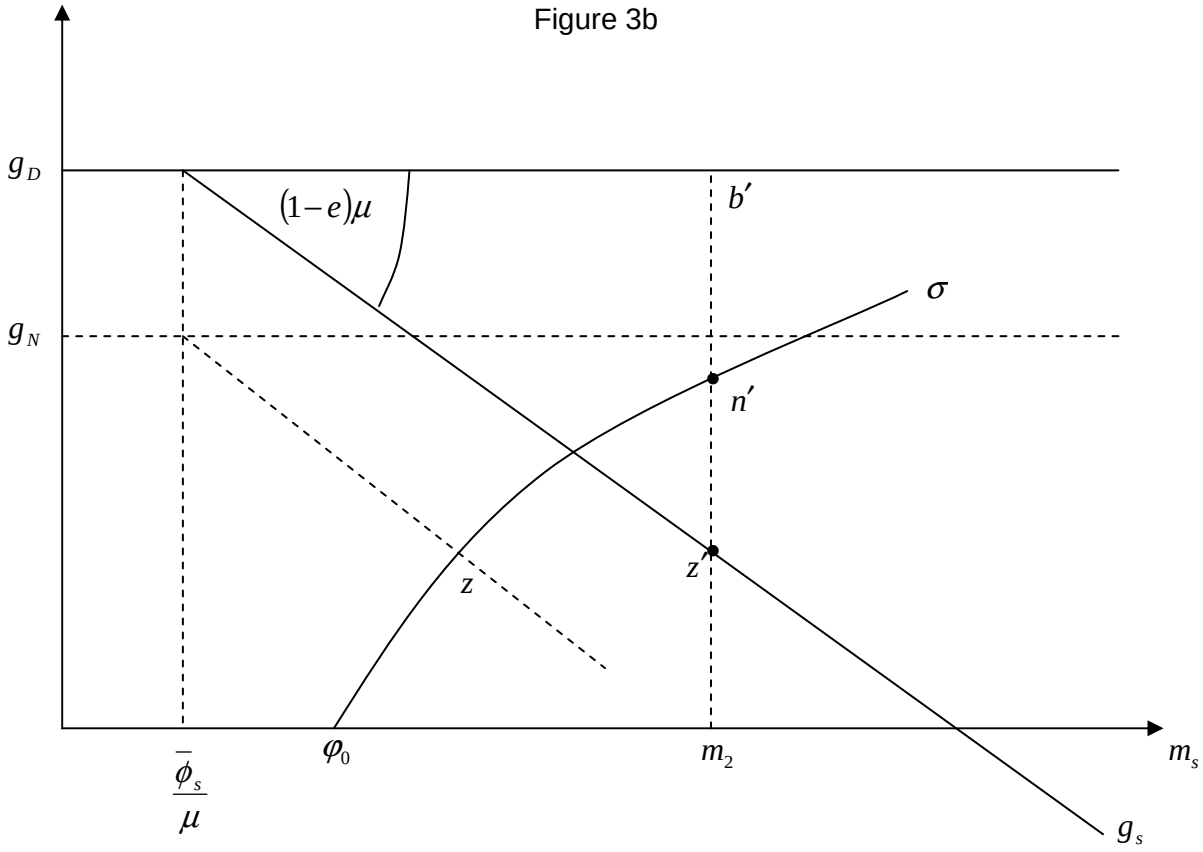


Figure 3b



The Entry and Exit Process