Dynamics of entry and exit of firms in the presence of entry adjustment costs[1]

Sanghamitra Das*, Satya P. Das

Indian Statistical Institute, 7 S.J.S. Sansanwal Marg, New Delhi 110016, India

Abstract

This paper presents a simple model of industry dynamics with entry adjustment costs. These costs imply a non-instantaneous adjustment path to the steady state, so that the model permits the short-run and long-run characterization of industry dynamics in a single framework. The type of the steady state (zero entry and exit or positive entry and exit with the co-existence of high-efficiency and low-efficiency firms) depends on the fixed cost of entry as well as entry adjustment costs. The short-run dynamics exhibits non-monotonicity if persistence in efficiency over time is not too high. The model is consistent with the empirical observations that the total number of firms and net entry follow non-monotonic paths and entry and exit are positively correlated over time.

Keywords: Industry evolution; Entry; Exit; Sunk cost

JEL classification: D4; D21; L1

1. Introduction

Many theoretical and empirical models of industry evolution focus on the movement of the number of firms and net entry over time. More recent empirical studies on industry dynamics indicate high rates of turnover in

* Corresponding author. Tel.: 91-11-651-4594; e-mail: dasm@isid.ernet.in.

[1] We are grateful to the referees whose comments led to major improvements of the paper. We thank Boyan Jovanovic, Ariel Pakes and Mark Roberts for helpful discussions. An earlier version of this paper was presented at the American Economic Association Meetings in Washington, DC, in January 1995. The usual disclaimer applies.

terms of entry as well as exit of firms in many countries. See Dunne et al. (1988) and Dunne and Roberts (1991) for patterns in the data for the US, Baldwin and Gorecki (1991) for Canada, Geroski's survey (1991, Chapter 2) for UK and other countries and Schwalbach (1991) for West Germany. These studies show that not only do entry and exit occur simultaneously but very interestingly they are positively correlated across industries at a point in time as well as over time within an industry. This rather curious phenomenon cannot be explained by conventional industry models assuming identical firms which have the feature that starting from a situation of zero entry and exit, a change in the environment either leads some firms to enter or exit but not both. In other words, they cannot explain the simultaneity of entry and exit, let alone the positive correlation between them.

Another empirical regularity observed is that the number of firms and net entry do not monotonically reach their steady state values (Gort and Klepper, 1982).

This paper presents a simple heterogeneous-firm dynamic model of an industry which generates these empirical regularities. The endogenous variables pertaining to the industry, such as entry, exit and the number of firms, follow a non-monotonic adjustment path to the steady state, and, furthermore, entry and exit are positively correlated *over time*. We also analyze the steady-state (long-run) and the short-run effects of demand and cost shocks. Along the steady states (where the rate of entry equals the rate of exit), a demand shock does not affect the rates of entry–exit but increases the total number of firms, while an increase in fixed costs increases the rate of entry–exit and decreases the number of firms. In contrast, in the short run a demand shift decreases both entry and exit, whereas a fixed-cost shift has ambiguous effects–entry and exit may change in the same or opposite direction. After the realization of the initial effects, they move in phase over time, irrespective of the source of initial perturbation. The model generates other testable hypotheses on co-movements as well.

Beginning with the pioneering work of Jovanovic (1982), attempts have been made recently to develop industry evolution models with heterogeneous firms that can explain the simultaneity of entry and exit, e.g. Lippman and Rumelt (1982), Jovanovic and Lach (1989), Lambson (1991, 1992), Hopenhayn (1992) and Ericson and Pakes (1989). Our model is similar to Hopenhayn in that firms are heterogeneous in their efficiency, which is subject to random shocks every period. The dynamics of our model is, however, quite different. Entry sunk costs in our model contain a fixed cost of entry (similar to Hopenhayn as well as Jovanovic, 1982, and Lippman and Rumelt, 1982), and, in addition, a variable component which is an increasing function of the number of entrants (as in Ericson and Pakes). The latter may be called entry adjustment costs. Such costs would arise due to resources needed for initial set-up that are imperfectly elastic in supply–

costs of resources which are wholly or partly specific to the industry. An example would be the cost of initial advertising to make consumers aware of a new brand. The greater the number of entrants the greater is the cost of product differentiation.

Not only are such entry adjustment costs realistic, they also prevent instantaneous adjustment of entry and exit. Unlike in Hopenhayn, the presence of these costs allows a framework within which to analyze short-run dynamics and patterns of co-movements. In particular, the co-movement between entry and exit over time becomes a non-trivial issue. This is unlike along the steady state where the rate of entry must equal the rate of exit and hence the positive correlation between them holds almost as an assumption rather than as an implication. The model developed in this paper is a simple example (less general than, for instance, Ericson and Pakes, and Hopenhayn) but the simplicity permits analytical characterization of the dynamics outside the steady state.

The model is laid out in Section 2. Steady-state and long-run effects of demand and cost shocks are analyzed in Section 3. The solution of the dynamic system, patterns of contemporaneous relationships along the dynamic path and the short-run effects of external shocks are examined in Section 4. The analysis of Section 4 includes the co-movement between entry and exit. The model generates other testable hypotheses as well. Concluding remarks are made in Section 5.

2. The model

We develop an industry evolution model of the simplest kind. It is assumed that the efficiency of a firm at a given point in time is a random outcome, though some persistency is allowed. We abstract from other factors that, in general, may have a bearing on the industry evolution process, such as effort, passive learning from experience or active learning through R&D, inventions and innovations, capital investment and mergers and acquisitions. While the existing literature on industry evolution mostly considers a perfectly competitive industry, we develop a model of imperfect competition. Within the class of non-competitive markets, entry being more common in industries with a large number of small firms rather than a small number of large firms, we assume a monopolistically competitive market rather than oligopoly[2].

[2] Londregan (1990) presents an interesting duopoly-rivalry model of industry evolution in which firms can enter, exit, re-enter and pre-empt each other. However, the simultaneity of entry and exit or the pattern of correlation between them is not the main issue.

2.1. Demand side

Each firm produces a unique brand (variety) of the same generic product. Hence, at any given time t, the number of firms operating, $n(t)$, equals the number of varieties available to consumers.

We assume a Dixit–Stiglitz utility function

$$\left[\sum_{i=1}^{n(t)} x_i(t)^{(\theta-1)/\theta}\right]^{\gamma\theta/(\theta-1)} [y(t)^{1-\gamma}], \quad 0<\gamma<1, \quad \theta>1,$$

where θ is the elasticity of substitution between any two brands, x_i the amount consumed of brand i and y the basket of other goods whose prices are fixed. The household maximizes utility subject to the budget

$$\sum_{i=1}^{n(t)} p_i(t)x_i(t) + y(t) \leq I,$$

where the price of y is normalized to unity and I is the total expenditure. The demand functions are explicitly solved as

$$x_i(t) = \left[\frac{I\gamma}{n(t)P(t)}\right]\left[\frac{p_i(t)}{P(t)}\right]^{-\theta} = \left[\frac{\theta\xi}{n(t)P(t)}\right]\left[\frac{p_i(t)}{P(t)}\right]^{-\theta}, \tag{1}$$

where $P(t)$ is a price index equal to

$$P(t) = \left[\frac{\sum_{i=1}^{n(t)} p_i(t)^{1-\theta}}{n(t)}\right]^{1/(1-\theta)}$$

and $\xi \equiv I\gamma/\theta$. There is no strategic interaction among firms. A firm i faces the demand curve (1) treating $n(t)$, $I\gamma$ and $P(t)$ as exogenous. Although the demand curve facing each firm is the same, each demand curve is for a distinct brand.

2.2. Cost side

The technology is described by the cost function: $C_i(x_i(t)) = \hat{c}_i x_i(t) + F$, where $\hat{c}_i$ is the marginal cost assumed independent of the output, and F is the fixed cost. Across the firms, $\hat{c}_i$s are random and take two possible values, $\hat{c}_H$ and $\hat{c}_L$ with $\hat{c}_H < \hat{c}_L$. Firms experiencing $\hat{c}_H$ are the high-efficiency (H-) firms and those experiencing $\hat{c}_L$ are the low-efficiency (L-) firms. This is the nature of heterogeneity across firms. Since the purpose is to highlight heterogeneity, it is presumed that $\hat{c}_L$ is sufficiently larger than $\hat{c}_H$ (see Condition R later). The $\hat{c}_i$s are assumed to follow a discrete first-order

Markov process (similar to that in Hopenhayn) with the transition probabilities given by

$$\begin{array}{ccc} & \hat{c}_H(t+1) & \hat{c}_L(t+1) \\ \hat{c}_H(t) & \rho & 1-\rho \\ \hat{c}_L(t) & 1-\rho & \rho\,, \end{array} \tag{2}$$

where $0.5 \leq \rho < 1$. $\rho = 0.5$ means that there is no persistence or firm-specific effects in the evolution process of the marginal cost; whether a firm at t was an H-firm or an L-firm, the odds of it being an H-firm or L-firm at time $t+1$ is 50–50. $\rho > 0.5$ means there are some firm-specific effects. By design, our model is meant to stress the role of randomness in the evolution of the marginal cost. It will thus be assumed that the difference between ρ and 0.5 be not too large (made precise in Condition R later). The entry process takes one period. Moreover, the entrants having no production experience before, face the odds ($\hat{c}_H$, $\hat{c}_L$: 0.5, 0.5) in the next period when production begins.

The transition probabilities given in (2) constitute the source of industry dynamics in this model. Based on (2) the incumbent firms as well as potential entrants calculate their "values"–the discounted sum of expected profits. Accordingly, the incumbent firms decide to stay on or leave. Potential entrants choose whether or not to enter the industry.

2.3. *Spot market equilibrium*

It is assumed that firms discover their type at the beginning of each period. A firm i of type j ($j =$ H, L) which stays maximizes profits, $\pi_{ij} = p_i(t)x_i(t) - \hat{c}_j x_i(t) - F$, subject to the demand curve it faces given in (1). The optimal pricing rule for firm i of type j is: $p_{ij}(t) = [\theta/(\theta-1)]\hat{c}_j$, where $\theta/(\theta-1)$ is the mark-up over the marginal cost. Firms of the same type exhibit identical behavior and hence from now onwards the subscript i will be suppressed.

Using the pricing rule, the reduced-form profit expressions are:

$$\pi_j(t) = \frac{\xi c_j(t)}{n_H(t)c_H(t) + n_L(t)c_L(t)} - F, \quad j = \text{H, L}, \quad c_j \equiv \hat{c}_j^{\,1-\theta}\,. \tag{3}$$

By definition, $c_H > c_L$ as $\theta > 1$. Clearly, $\pi_H(t) > \pi_L(t)$, for any $n_H(t)$ and $n_L(t)$.

2.4. *Entry and exit processes and the value functions*

During any period, events follow the following sequence:

Realization of marginal cost	→	Some firms may exit	→	Remaining firms produce and sell. Some firms may enter

Thus, at the beginning of the period, firms realize their type. Then some firms may exit. Remaining firms operate and at the same time new firms may start to enter. As said earlier, the entry process takes one period.

We assume that the entry (sunk) cost function takes the form:

$$k + mn_{\mathrm{W}}(t), \quad k, m > 0 . \tag{4}$$

k is the fixed costs of entry such as any entry fee or plant set-up costs. $mn_{\mathrm{W}}(t)$ represents the adjustment costs of entry, which are likely to exist in a differentiated product industry as discussed in Section 1; these costs prevent instantaneous adjustment of the industry to the steady state and are critical to our analysis.

If there is entry, the zero-profit condition for entrants is given by the value of the new entrant, $v_{\mathrm{W}}(t)$, being equal to the total sunk cost of entry:

$$v_{\mathrm{W}}(t) = k + mn_{\mathrm{W}}(t) . \tag{5}$$

In view of the transition probability matrix (2), the values of entering firms as well as the existing firms are, by definition,

$$v_{\mathrm{H}}(t) = \pi_{\mathrm{H}}(t) + \beta[\rho \max(s, v_{\mathrm{H}}(t+1)) + (1-\rho)\max(s, v_{\mathrm{L}}(t+1))] \tag{6a}$$

$$v_{\mathrm{L}}(t) = \pi_{\mathrm{L}}(t) + \beta[\rho \max(s, v_{\mathrm{L}}(t+1)) + (1-\rho)\max(s, v_{\mathrm{H}}(t+1))] \tag{6b}$$

$$v_{\mathrm{W}}(t) = \beta[0.5 \max(s, v_{\mathrm{H}}(t+1)) + 0.5 \max(s, v_{\mathrm{L}}(t+1))] , \tag{6c}$$

where β is the discount factor and $s > 0$ is the opportunity cost of being in the industry. s is assumed the same for an H-firm or an L-firm.

2.5. *Industry dynamics*

There are four possible paths along which the industry may evolve: (a) no entry or exit, (b) exit but no entry, (c) entry but no exit and (d) entry and exit. We will, however, concentrate on dynamics in the neighborhood of the steady state. Since the total number of firms along the steady state is given, there is either no entry–exit or positive entry–exit. Thus (a) and (d) are only relevant. The dynamics of these cases are outlined below.[3]

[3] A previous version of this paper contains the description of cases (b) and (c) as well.

2.5.1. No entry or exit

Consider a time interval Δ. No entry over this interval means that entry is an unprofitable proposition for any number of entrants, i.e.

$$v_W(t) < k, \quad t \in \Delta. \tag{7}$$

No exit implies

$$v_L(t) > s, \quad t \in \Delta. \tag{8}$$

In view of (8), the value functions (6a)–(6c) reduce to

$$v_H(t) = \pi_H(t) + \beta[\rho v_H(t+1) + (1-\rho)v_L(t+1)] \tag{9a}$$

$$v_L(t) = \pi_L(t) + \beta[\rho v_L(t+1) + (1-\rho)v_H(t+1)] \tag{9b}$$

$$v_W(t) = \frac{\beta}{2}[v_H(t+1) + v_L(t+1)]. \tag{9c}$$

Assuming that the law of large number applies and hence no aggregate uncertainty, the composition of firms evolves non-stochastically in accordance with the transition probabilities:

$$\begin{aligned} n_H(t+1) &= \rho n_H(t) + (1-\rho)n_L(t) \\ n_L(t+1) &= \rho n_L(t) + (1-\rho)n_H(t). \end{aligned} \tag{10}$$

Since there is no entry and exit, the total number of firms remains unchanged, say, equal to $\bar{n}$. Then $n_H(t+1) = (2\rho - 1)n_H(t) + (1-\rho)\bar{n}$. Thus, if $\rho = 0.5$, $n_H(t)$ and $n_L(t)$ are constant, equal to $\bar{n}/2$ each. If $\rho > 0.5$, $n_H(t)$ and $n_L(t)$ converge monotonically to $\bar{n}/2$. Although possible, this is the most uninteresting outcome given that our objective is to study entry and exit.

2.5.2. Entry and exit

This is the case of simultaneous entry and exit. As long as there is entry it follows that none of the firms that find themselves in the beginning of a period as the high-efficiency type exit. Because it would mean that in equilibrium an H-firm earns s, hence an entering firm can at best earn the opportunity cost and so there would be no incentive to enter and incur the sunk cost of entry, i.e. there will be no entry. This implies that (positive) entry is consistent with some or all of L-firms but none of H-firms quit.[4]

[4] We are grateful to a referee who brought our attention to the case where all L-firms quit. Recognition of this possibility has corrected an error in the steady-state analysis contained in an earlier version of this paper.

2.5.2.1. Some L-firms quit

Since no H-firms quit and some L-firms do,

$$v_H(t) > s, \quad v_L(t) = s \,. \tag{11}$$

The zero-profit condition for the entrants holds, i.e.

$$v_W(t) = k + mn_W(t) \,. \tag{5}$$

In view of (6a)–(6c), $v_L(t) = s$ implies

$$v_H(t) = \pi_H(t) + \beta[\rho v_H(t+1) + (1-\rho)s] \tag{12a}$$

$$s = \pi_L(t) + \beta[(1-\rho)v_H(t+1) + \rho s] \tag{12b}$$

$$v_W(t) = \frac{\beta}{2}[v_H(t+1) + s] \,. \tag{12c}$$

Note that (12b) is the exit equilibrium condition. The number of H-firms evolves according to

$$n_H(t+1) = \rho n_H(t) + (1-\rho)n_L(t) + \frac{1}{2}n_W(t) \,. \tag{13}$$

The rate of exit or the number of quits, say $q(t)$, has a complicated expression, equal to the difference between (a) the number of L-firms at the beginning of the period, equal to $(1-\rho)n_H(t-1) + \rho n_L(t-1) + (1/2)n_W(t-1)$ and (b) the number of existing L-firms in the period, equal to $n_L(t)$. In other words,

$$q(t) = (1-\rho)n_H(t-1) + \rho n_H(t-1) + \frac{1}{2}n_W(t-1) - n_L(t) \,. \tag{14}$$

Eqs. (5), (12a)–(12c), (13) and (14) govern the industry dynamics. They are six equations in six variables: $v_H(.)$, $v_W(.)$, $n_H(.)$, $n_L(.)$, $n_W(.)$ and $q(.)$.

The dimension of the dynamic system can be reduced. Using (3) (the expression for instantaneous profit) the exit equilibrium condition (12b) is equivalent to

$$n_L(t) = \frac{\xi}{F + (1-\beta\rho)s - \beta(1-\rho)v_H(t+1)} - \frac{c_H}{c_L}n_H(t) \,. \tag{15}$$

(12c) and (5) can be combined to eliminate $v_W(t)$ and yield

$$n_W(t) = \frac{\beta}{2m}[v_H(t+1) + s] - \frac{k}{m} \,. \tag{16}$$

In what follows, the dynamics will be summarized in the $(n_H(.), v_H(t))$ space. Hence we need the expressions for the rates of change in $n_L(t)$ and $n_W(t)$ with respect to $n_H(t)$ and $v_H(t+1)$, although the latter variables are also endogenous to the model.

Eqs. (15) and (16) express $n_L(t)$ and $n_W(t)$ in terms of $n_H(t)$ and $v_H(t+1)$ with

$$\frac{\partial n_L(t)}{\partial n_H(t)} = -\frac{c_H}{c_L} < 0 \tag{17a}$$

$$\frac{\partial n_L(t)}{\partial v_H(t+1)} = \frac{\beta(1-\rho)\xi}{[F + (1-\beta\rho)s - \beta(1-\rho)v_H^*]^2} > 0 \tag{17b}$$

$$\frac{\partial n_W(t)}{\partial n_H(t)} = 0 \tag{17c}$$

$$\frac{\partial n_W(t)}{\partial v_H(t+1)} = \frac{\beta}{2m} > 0\,. \tag{17d}$$

$\partial n_L(t)/\partial(v_H(t+1))$ above is a local property; it holds in the neighborhood of the steady state along which $v_H(t) = v_H^*$. Other expressions are global.

Substituting (15) and (16) into (13) implicitly defines (18) below and its derivatives:

$$n_H(t+1) = h[n_H(t), v_H(t+1)], \quad \text{where} \tag{18}$$

$$h_n = 2\rho - 1 - (1-\rho)\left(\frac{c_H}{c_L} - 1\right)$$

$$h_v = (1-\rho)\frac{\partial n_L}{\partial v_H} + \frac{\beta}{4m} > 0\,.$$

Next, if we substitute (3) into Eqs. (12a) and (12b) and eliminate $\pi_H(t)$ and $\pi_L(t)$,

$$\begin{aligned} v_H(t+1) &= -\frac{c_L}{\beta[(1-\rho)c_H - \rho c_L]} v_H(t) \\ &\quad + \frac{(c_H - c_L)F + [c_H(1-\beta\rho) + c_L\beta(1-\rho)]s}{\beta[(1-\rho)c_H - \rho c_L]} \\ &\equiv g[v_H(t)] \end{aligned} \tag{19}$$

$$g_v = -\frac{c_L}{\beta[(1-\rho)c_H - \rho c_L]}\,.$$

Eqs. (18) and (19) constitute two first-order difference equations in two variables, $n_H(t)$ and $v_H(t)$. There is one initial condition, $n_H(0) = n_H^0$. Other variables are solved in terms of the solutions of $n_H(t)$ and $v_H(t)$.

2.5.2.2. All L-firms quit

It is possible that all L-firms quit. Hence at any given point in time all operating firms in the industry are H-firms. The zero-profit condition (5) holds. The value of an entering firm equals

$$v_W(t) = \frac{\beta}{2}[v_H(t+1) + s], \tag{12c}$$

as in the earlier case. Moreover,

$$v_H(t) = \pi_H(t)|_{n_L(t)\to 0} + \beta[\rho v_H(t+1) + (1-\rho)s]. \tag{20}$$

Since $n_L(t) = 0$, $n_H(t)$ evolves according to

$$n_H(t+1) = \rho n_H(t) + \frac{1}{2} n_W(t). \tag{21}$$

For all L-firms to exit, $v_L(t) \leq s$ must hold as $n_L(t) \to 0$, i.e.

$$\pi_L(t)|_{n_L(t)\to 0} + \beta[(1-\rho)v_H(t+1) + \rho s] \leq s. \tag{22}$$

Eqs. (5), (12c), (20) and (21) contain four variables: $v_H(.)$, $v_W(.)$, $n_H(.)$ and $n_W(.)$. From our viewpoint, this is rather an uninteresting case, although more interesting than the zero entry–exit case. While there is simultaneity of entry and exit, the operating firms (those who stay and hence are "observed") are all homogeneous. We will therefore skip further characterization of the dynamics of this case.

2.6. A recapitulation

It seems useful at this point to briefly recapitulate various elements of the model. First, the efficiency-heterogeneity is assumed to be an outcome of chance mechanism only, allowing for some persistence. In general of course, efficiency is influenced by a host of other factors that are not captured in the model, such as capital accumulation, effort, learning by doing, R&D investment, development of outside knowledge and the like. However, the randomness of efficiency captured in the model is something inherent in any industry, not industry-specific. Second, with respect to entry and exit, hysteresis effects of parametric changes are quite possible in the model at the individual firm level not only due to the sunk cost of entry (see, for example, Baldwin, 1988; Dixit, 1989) but also because a firm's efficiency may change over time. Third, there is individual-firm uncertainty but no aggregate uncertainty. Fourth, there is perfect foresight. Fifth, unlike standard dynamic models where individual behavior is specified and then aggregated, here the identity of the individual firm in terms of efficiency changes over time. At each point in time, potential entrants decide whether or not to enter. Incumbents decide whether or not to stay and if they stay how much to produce and what price to charge.

3. Steady-state analysis

The size as well as the composition of the industry remains unchanged along the steady state. Hence the rate of exit must be equal to the rate of entry. There are two cases accordingly: positive entry–exit and zero entry–exit. There are two subcases to the former: some of the L-firms quit and all L-firms quit. We turn to each of these cases.

3.1. Positive entry and exit, and some of the L-firms exit

In the steady state the value remains unchanged, i.e. $v_H(t) = v_H(t+1) = v^*$. Eq. (19) yields a closed-form solution for the long-run value of an H-firm:

$$v^* = \frac{(c_H - c_L)F + [c_H(1-\beta\rho) + c_L\beta(1-\rho)]s}{\beta(1-\rho)c_H + (1-\beta\rho)c_L} . \tag{23a}$$

It is interesting to note that v^* is unaffected by a change in ξ (the market size) and increases with F. This follows from two factors: (a) the exit equilibrium which implies that the value of an L-firm remains equal to s irrespectively of ξ and F and (b) firm heterogeneity. v^* is unaffected by a change in ξ since a change in ξ leads to the same proportionate increase in the operating profits whether it is an H- or an L-firm. An increase in F tends to reduce $\pi_H(t)$ less than it does $\pi_L(t)$ relative to the respective operating profits. The value of an L-firm remains unchanged ($=s$), in equilibrium the value of an H-firm increases. Further understanding is gained from analyzing (12a) and (12b) which describe the definition of the value and the exit equilibrium. Note that $\pi_H(t)$ and $\pi_L(t)$ are related by

$$\pi_H(t) = \frac{c_H}{c_L}\pi_L(t) + \left(\frac{c_H}{c_L} - 1\right)F .$$

Substituting this into (12a), the steady-state versions of (12a) and (12b) can be expressed as

$$(1-\beta\rho)v^* = \frac{c_H}{c_L}\pi_L^* + \left(\frac{c_H}{c_L} - 1\right)F \tag{23b}$$

$$(1-\beta\rho)s = \pi_L^* + \beta(1-\rho)v^* . \tag{23c}$$

These are two equations in two variables, v^* and π_L^*, and independent of ξ. Thus $dv^*/d\xi = 0$. From (23c), as F increases, either (i) π_L^* increases and

v^* decreases or (ii) π_L^* decreases and v^* increases. But (i) is inconsistent with (23b); thus (ii) holds.

The solutions of (23b) and (23c) yields (23a) as well as

$$\pi_L^* \equiv \phi = \frac{[(1-\beta\rho)^2 - \beta^2(1-\rho)^2]c_L s - \beta(1-\rho)(c_H - c_L)F}{\beta(1-\rho)c_H + (1-\beta\rho)c_L}. \tag{24}$$

Given v^*, the rate of entry and exit (the rate of entry equals the rate of exit) is determined from (16):[5]

$$n_W^* = \frac{\beta(v^*+s)}{2m} - \frac{k}{m}. \tag{25}$$

In the steady state, (13) reduces to

$$n_H^* = n_L^* + \frac{n_W^*}{2-2\rho}. \tag{26}$$

Substituting (25) and (26) into the steady-state version of (15) solves n_L^*:

$$n_L^* = \frac{1}{c_H + c_L}\left[\frac{c_L \xi}{F+\phi} - \frac{c_H[\beta(v^*+s) - 2k]}{4m(1-\rho)}\right]. \tag{27}$$

Given n_W^* and n_L^*, Eq. (26) determines n_H^*.

A steady state with positive entry and exit implies that these solutions of n_W^* and n_L^* must be positive. From the respective solution expressions:

$$n_W^* > 0 \Leftrightarrow k < \frac{\beta(v^*+s)}{2} \tag{28a}$$

$$n_L^* > 0 \Leftrightarrow k > \frac{\beta(v^*+s)}{2} - \frac{2(1-\rho)mc_L}{(F+\phi)c_H}\xi. \tag{28b}$$

This pair of inequalities is illustrated in the (k, ξ) space in Fig. 1. $n_W^* > 0$ is satisfied below the AA line. $n_L^* > 0$ is met to the right of the downward sloping BB line. Thus the positive entry–exit steady state with some of L-firms exiting occurs in the region B.

3.2. Positive entry and exit, and all L-firms exit

Recall that Eqs. (5), (12c), (20) and (21) describe the dynamics in this case. The steady-state versions of these equations are:

$$v_W^* = k + mn_W^* \tag{29a}$$

$$v_W^* = \frac{\beta}{2}(v^*+s) \tag{29b}$$

[5] From (14), $q^* = (1-\rho)n_H^* + \rho n_L^* + (1/2)n_W^* - n_L^*$. Substituting (26) into it yields $q^* = n_W^*$.

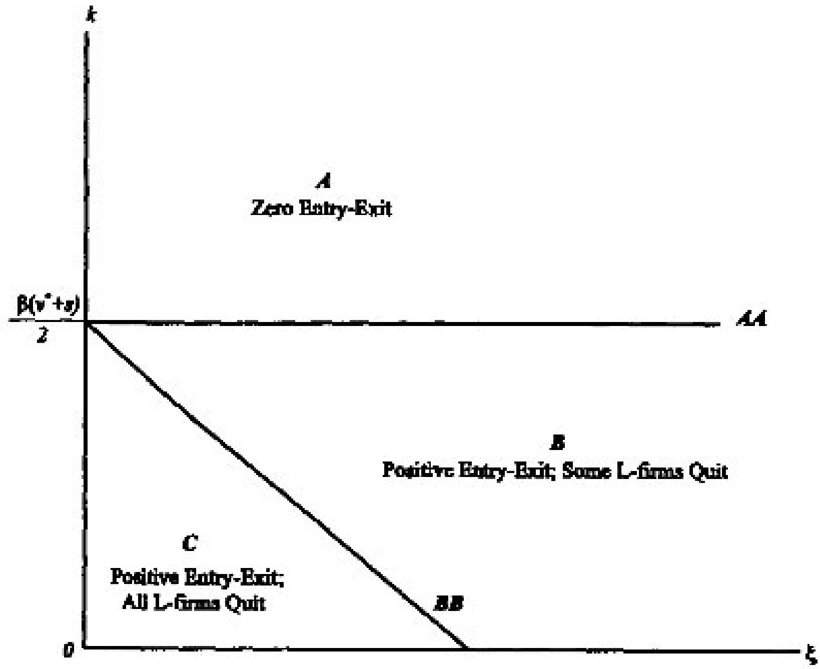

Fig. 1. Steady state.

$$v^* = \frac{\xi}{n_H} - F + \beta[\rho v^* + (1-\rho)s] \tag{29c}$$

$$n_H^* = \rho n_H^* + \frac{n_W^*}{2}. \tag{29d}$$

Eliminating v_W^*, v^* and n_W^*, these equations imply

$$k + 2m(1-\rho)n_H^* - \frac{\beta\xi}{2(1-\beta\rho)n_H^*} = \frac{s\beta(1+\beta-2\beta\rho)}{2(1-\beta\rho)} - \frac{F\beta}{2(1-\beta\rho)}, \tag{30}$$

which determines n_H^*. Eq. (30) determines a unique value of n_H^* in R_+. Moreover, it implicitly defines $n_H^* \equiv n(k, \xi)$ such that an increase in k and an increase in ξ, respectively, decrease and increase n_H^*, as expected.

Since all L-firms quit, $v_L^* < s$ as $n_L^* \to 0$. In view of (3), (29c) and (30), the steady-state version of (22) is equivalent to

$$\left[\frac{c_L}{c_H n(k,\xi)} + \frac{\beta(1-\rho)}{(1-\beta\rho)n(k,\xi)}\right]\xi \leq \frac{(1-\beta)(1+\beta-2\beta\rho)}{1-\beta\rho}s + \frac{1+\beta-2\beta\rho}{1-\beta\rho}F. \quad (31)$$

It is straightforward to check that in the (k, ξ) space, this inequality is satisfied in region C indicated in Fig. 1.[6] We now turn to the remaining case of no entry–exit.

3.3. *No entry and no exit*

The steady state in this case is characterized by

$$v_W^* < k \quad (32a)$$

$$v_L^* \geq s \quad (32b)$$

$$(1-\beta\rho)v^* = \pi_H^* + \beta(1-\rho)v_L^* \quad (32c)$$

$$(1-\beta\rho)v_L^* = \pi_L^* + \beta(1-\rho)v^* \quad (32d)$$

$$v_W^* = \frac{\beta}{2}(v^* + v_L^*) \quad (32e)$$

$$n_H^* = n_L^*, \quad (32f)$$

which correspond to (7), (8), (9a)–(9c) and (10). Substituting (32f) into (3),

$$\pi_j^* = \frac{c_j \xi}{n_H^*(c_H + c_L)} - F, \quad j = H, L. \quad (33)$$

Using (33) we can now solve for v^*, v_L^* and v_W^* in terms of n_H^* from (32c)–(32e). The inequalities (32a) and (32b) can then be equivalently stated as

$$n_H^* > \frac{\beta\xi}{2[(1-\beta)k + F\beta]} \quad (34a)$$

$$n_H^* \leq \frac{\xi[\beta(1-\rho)c_H + (1-\beta\rho)c_L]}{(1-\beta-2\beta\rho)(c_H + c_L)[F + (1-\beta)s]} = \frac{c_L \xi}{(c_H + c_L)(F + \phi)}. \quad (34b)$$

These inequalities must be satisfied at the steady state if there is no entry or exit. The inequality (34a) says that there is no entry when the fixed cost

[6] The equality version of (31) defines the *BB* line.

of entry is too high relative to the value of entry or the number of firms is too large such that the value of entry falls short of the fixed cost of entry. The inequality (34b) states that there will be no exits if the market size, indicated by ξ, is sufficiently large or the total number of firms is sufficiently small or both so that the instantaneous profits and the value of being an L-firm is large enough.

Using the expression for ϕ, (34a) and (34b) imply

$$k > \frac{\beta}{2}(v^* + s)$$

This is readily identified in Fig. 1–region A above the line AA.

3.4. Propositions

Different types of steady states are characterized in Fig. 1. Hence

Proposition 1. There exists steady state with no entry–exit, positive entry–exit with co-existence of high- and low-efficiency firms and positive entry–exit with only high-efficiency firms remaining in the industry according as (k, ξ) *belong to region A, B or C.*

Remarks:

- Proposition 1 ties with Hopenhayn's characterization of the steady state given in his Theorem 3. Similar to Hopenhayn, positive or zero entry–exit steady state obtains as the fixed sunk cost of entry is below or above a critical level.
- Our model, however, provides further characterization of the positive entry–exit steady state. Given that k is low enough so that there is a positive entry–exit steady state, high- and low-efficiency firms co-exist if market size is large and only high-efficiency firms operate otherwise.
- The entry adjustment cost m and the persistence parameter ρ have implications for steady state too. The higher m is, the steeper the BB line and the greater the likelihood of positive entry–exit steady state with co-existence of both types of firms. The higher ρ is, the flatter the BB line and less is the likelihood of such a steady state.

Clearly, region B is the most interesting case: there is positive entry and exit and heterogeneous firms co-exist and operate in the industry. From now onwards it will be presumed that parametric configurations belong to this region.

The model permits a number of interesting comparative static analyses. We will limit ourselves to (permanent) changes in demand (ξ) and fixed costs (F). Assuming that parametric changes are not too large to displace the steady state from region B to region A or C, the following comparative statics are immediate from Eqs. (23)–(27):

$$\frac{dn_W^*}{d\xi}=0;\quad \frac{dn_L^*}{d\xi}>0;\quad \frac{dn_H^*}{d\xi}>0;\quad \frac{dn^*}{d\xi}>0 \tag{35}$$

$$\frac{dn_W^*}{dF}>0;\quad \frac{dn_L^*}{dF}<0;\quad \frac{dn_H^*}{dF}\gtrless 0;\quad \frac{dn^*}{dF}<0\,, \tag{36}$$

where n^* denotes the total number of firms. Note that along the steady state a change in n_W^* represents a change in the rate of entry as well as the rate of exit. Thus

Proposition 2. Along the steady states: (a) A permanent increase in demand does not affect the rates of entry and exit. It supports a larger number of H- and L-firms and hence a larger number of firms in total. (b) A permanent increase in fixed costs leads to an increase in the rates of entry and exit. The number of L-firms declines, the number of H-firms may increase or decrease and the total number of firms decreases.

4. Analysis of the transition path

We now analyze the industry evolution process along the transition path to the steady state. We will be concerned with local dynamics. Suppose an initial steady state (in region B of Fig. 1) is disturbed by some shock and in Period 0, $n_H(0)\equiv n_H^0 \neq n_H^*$. How do entry, exit, composition and total size of the industry change over time?

We turn to the difference-equation system (18)–(19). Note that this is a recursive system. Hence the eigen roots are equal to h_n and g_v. The existence of a unique perfect foresight path requires that the modulus of h_n or g_v, but not both, be less than one. However, from (18) and (19), the magnitude or sign of h_n and g_v is not apparent. The Appendix discusses how their signs and magnitudes change as ρ varies from 0.5 to 1. It is shown that if a "regularity condition" (Condition R below) is satisfied, $|h_n|>1$ and $|g_v|<1$. Hence a perfect foresight path exists with g_v as the stable root of the system. Moreover, $g_v<0$.

$$\text{Condition R:}\quad \frac{c_H}{c_L}>\max\left(\frac{2+\beta}{\beta},\ \frac{1+\rho}{1-\rho}\right).$$

Remarks:

• Condition R essentially means two things. (a) Heterogeneity among the two types of firms should be sufficiently large, i.e. $\hat{c}_L$ (c_L) is sufficiently larger (smaller) than $\hat{c}_H$ (c_H). (b) The difference between ρ and 0.5 should not be too large.

Both requirements are consistent with the spirit of our model. Firm heterogeneity is the hallmark of industry evolution models in general including ours. Our model is designed to highlight the randomness in the evolution of cost function facing the firms. It is befitting, therefore, that persistence in costs (the difference between ρ and 0.5) not be too high.

Moreover, as seen earlier, the greater the value of ρ, the less likely it is that a steady state with positive entry–exit and co-existence of high- and low-efficiency firms will exist. Since our main purpose is to study entry and exit, we are motivated to assume that $\rho - 0.5$ is not too large.

• A reasonable lower bound for the discount rate β is 0.9. At this value, $c_H/c_L > (2+\beta)/\beta$ reduces to $c_H/c_L > 3.22$. However, in order for this inequality to be met, it is not necessary that the ratio of marginal costs, $\hat{c}_L/\hat{c}_H$ be at least 3.22. It is because $c_i = \hat{c}_i^{1-\theta}$ where $\theta > 1$, so a given ratio of $\hat{c}_L$ and $\hat{c}_H$ implies a larger ratio of c_H and c_L. For example, let $\hat{c}_H = 1$, $\hat{c}_L = 1.5$ and $\theta = 4$. Then $c_H/c_L = 3.375 > 3.22$. For any given $\hat{c}_H$ and $\hat{c}_L$, c_H/c_L is larger, larger the value of θ. In a monopolistically competitive market with a large number of substitute varieties, the elasticity of substitution, θ, can indeed be very high. Thus a small difference between $\hat{c}_L$ and $\hat{c}_H$ can ensure $c_H/c_L > (2+\beta)/\beta$.

The inequality $c_H/c_L > (1+\rho)/(1-\rho)$ requires that c_H/c_L be large and ρ be small relative to each other. In the above example, this inequality is equivalent to $\rho < 0.54$. But if θ or c_H/c_L were larger, the range of ρ for which this inequality is satisfied will be larger. For example, if $\hat{c}_H = 1$, $\hat{c}_L = 2$ and $\theta = 4$, then $c_H/c_L = 8$ and $\rho < 0.77$ will satisfy it.

The upshot of this remark is that Condition R is not too stringent in the context of the model.

Proposition 3. Under Condition R, a unique perfect foresight path exists with the associated eigen root equal to g_v.

In what follows we assume that Condition R is met. Denoting the deviation from the respective steady state by a tilde, the local solutions are then

$$\tilde{n}_H(t) = (n_H^0 - n_H^*)(g_v)^t \tag{37a}$$

$$\tilde{v}_H(t) = \frac{g_v - h_n}{h_v}(n_H^0 - n_H^*)(g_v)^{t-1}\,. \tag{37b}$$

$h_v > 0$ unambiguously, and under Condition R, $-1 < g_v < 0$ and $h_n < -1$. Thus

$$\frac{g_v - h_n}{h_v} > 0$$

Using (37a) and (37b) and differentiating (15) and (16)

$$\begin{aligned}\tilde{n}_{\rm L}(t) &= \left[\frac{\partial n_{\rm L}}{\partial n_{\rm H}} + \frac{\partial n_{\rm L}}{\partial v_{\rm H}}\left(\frac{g_v - h_n}{h_v}\right)\right](n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t \\ &= -\left[1 + \frac{\partial n_{\rm L}}{\partial v_{\rm H}}\left(\frac{2\rho - 1 - g_v}{h_v}\right) + \frac{\beta}{4mh_v}\left(\frac{c_{\rm H}}{c_{\rm L}} - 1\right)\right](n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t \\ &\equiv -a_1(n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t, \quad a_1 > 0 \end{aligned} \tag{38a}$$

$$\begin{aligned}\tilde{n}_{\rm W}(t) &= \frac{\partial n_{\rm W}}{\partial v_{\rm H}}\left(\frac{g_v - h_n}{h_v}\right)(n_{\rm h}^0 - n_{\rm H}^*)(g_v)^t = \frac{\beta}{2m}\left(\frac{g_v - h_n}{h_v}\right)(n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t \\ &\equiv a_2(n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t, \quad a_2 > 0 . \end{aligned} \tag{38b}$$

Thus

$$\begin{aligned}\tilde{n}(t) = \tilde{n}_{\rm H}(t) + \tilde{n}_{\rm L}(t) &= -\left[\frac{\beta}{4mh_v} + \frac{\partial n_{\rm L}}{\partial v_{\rm H}}\left(\frac{2\rho - 1 - g_v}{h_v}\right)\right](n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t \\ &\equiv -a_3(n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t, \quad a_3 > 0 . \end{aligned} \tag{39}$$

Differentiating (14) (the solution for the number of exits) and after appropriate substitutions,

$$\begin{aligned}\tilde{q}(t) &= \left[\frac{1-\rho}{g_v} + \left(\frac{\rho}{g_v} - 1\right)\left\{\frac{\partial n_{\rm L}}{\partial n_{\rm H}} + \frac{\partial n_{\rm L}}{\partial v_{\rm H}}\left(\frac{g_v - h_n}{h_v}\right)\right\}\right. \\ &\quad \left. + \frac{1}{2g_v}\frac{\partial n_{\rm W}}{\partial v_{\rm H}}\left(\frac{g_v - h_n}{h_v}\right)\right](n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t \\ &= \left[\left(1 - \frac{2\rho - 1}{g_v}\right)\left(1 + \frac{\beta(c_{\rm H}/c_{\rm L} - 1)}{4mh_v}\right) + \left(\frac{2\rho - 1 - g_v}{h_v}\right)\right. \\ &\quad \left. \times \left\{\left(1 - \frac{\rho}{g_v}\right)\frac{\partial n_{\rm L}}{\partial v_{\rm H}} - \frac{\beta}{4mg_v}\right\}\right](n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t \\ &\equiv a_4(n_{\rm H}^0 - n_{\rm H}^*)(g_v)^t, \quad a_4 > 0 . \end{aligned} \tag{40}$$

The coefficient of none of the above is zero, and in particular, $g_v < 0$. Hence the model exhibits the property that

Proposition 4. The value of an H-firm, the number and composition of firms, entry and exit and net entry follow non-monotonically converging adjustment time paths to the steady state.

The non-monotonic path for the industry stems from the result that the value of being an H-firm follows a non-monotonic path. Non-monotonicity of the value, in turn, arises from firm-heterogeneity together with the absence of strong persistence ($\rho \approx 0.5$). This is seen from (12a) and (12b) that describe, respectively, the definition of the value and the exit equilibrium. When $\rho \approx 0.5$, these equations reduce to

$$v_{\mathrm{H}}(t) = \pi_{\mathrm{H}}(t) + \frac{\beta}{2}[v_{\mathrm{H}}(t+1) + s] \tag{41a}$$

$$s = \pi_{\mathrm{L}}(t) + \frac{\beta}{2}[v_{\mathrm{H}}(t+1) + s]\,. \tag{41b}$$

It will be argued that an increase in $v_{\mathrm{H}}(t+1)$ must be associated with a decrease in $v_{\mathrm{H}}(t)$ in order for both equations to be satisfied. Suppose $v_{\mathrm{H}}(t+1)$ increases. Then (41b) implies that $\pi_{\mathrm{L}}(t)$ must decrease for the exit equilibrium condition to be satisfied. From (3), $\pi_{\mathrm{H}}(t)$ and $\pi_{\mathrm{L}}(t)$ are related by $\pi_{\mathrm{H}}(t) + F = (c_{\mathrm{H}}/c_{\mathrm{L}})[\pi_{\mathrm{L}}(t) + F]$. Thus $\delta\pi_{\mathrm{H}}(t) = (c_{\mathrm{H}}/c_{\mathrm{L}})\delta\pi_{\mathrm{L}}(t)$, implying that a given decrease in $\pi_{\mathrm{L}}(t)$ is associated with a larger decrease in $\pi_{\mathrm{H}}(t)$. Hence the right-hand side of (41a) decreases as $v_{\mathrm{H}}(t+1)$ increases. Thus $v_{\mathrm{H}}(t)$ must decrease.

The non-monotonicity of $n_{\mathrm{H}}(t)$, $n_{\mathrm{L}}(t)$, $n_{\mathrm{W}}(t)$ and $q(t)$ implies the total number of firms and net entry follow non-monotonic paths as well–which agrees with empirical evidence as reported in Gort and Klepper.

Co-movements of different variables are now apparent from (37a), (37b), (38a), (38b), (39) and (40). In particular, the nature of co-movement between entry and exit is seen from dividing (38b) by (40):

$$\frac{\tilde{n}_{\mathrm{W}}(t)}{\tilde{q}(t)} = \frac{a_2}{a_4} > 0\,, \tag{42}$$

that is,

Proposition 5. Along the adjustment path, the rates of entry and exit are positively related.

There is no "quick" explanation of this positive correlation. Suppose that the industry is initially in the steady state (with positive entry and exit) and there is a permanent increase in market demand, represented by an increase in ξ. The industry is thrown out of steady state. The new steady state number of high-efficiency firms is higher (see Proposition 2). In accordance with Proposition 4, the number of H- and L-firms as well as entry and exit follow non-monotonic paths toward the new steady state. Suppose during the adjustment path, $n_{\mathrm{H}}(t)$ increases from one period to the next, say from 1 to 2, that is, $n_{\mathrm{H}}(2) > n_{\mathrm{H}}(1)$. This tends to reduce current profits of both types of firms. In particular, the decrease in the current profit of an L-firm tends to decrease the current value of being an L-firm and hence forces more exits in Period 2 compared with Period 1. This is a *crowding-out* or a displacement effect on exit.

Entry in Period 2 depends upon the future value, $v_H(3)$, as entry takes one period. There are two effects on the future value. First, as seen in (12b), the decrease in the current profit of an L-firm tends to increase the future value of being an H-firm because in equilibrium the current value of being an L-firm, $\pi_L(2) + \beta[(1-\rho)v_H(3) + \rho s]$, must remain unchanged, equal to s. Second, the decrease in the current profit of an H-firm, at any given current value, $v_H(t)$, also tends to increase the future value (see (12a)). Rational, potential entrants disregard the effect of larger number of firms on the current profit (of being either an H- or an L-firm) but anticipate the increase in the future value. This encourages more entry. It may be called the *future-value effect*.

Thus, compared with Period 1, there is more entry and exit in Period 2. We have discussed what happens to entry and exit when $n_H(t)$ increases over time. The opposite happens and entry and exit rates both decline when $n_H(t)$ decreases over time. Although we have illustrated the case of demand shock, *any* shock that displaces the steady state is followed by a positive association between entry and exit over time after the initial period. In summary then, there are two effects that generate the positive relationship: a crowding-out effect on exit, which is straightforward, and a future-value effect on entry, which is more subtle. Hence the model is consistent with the positive correlation between entry and exit over time within an industry observed in empirical studies.

In the process, the transition equations (37a), (37b), (38a), (38b), (39) and (40) generate other testable hypotheses on contemporaneous correlations (not causations):

$$\frac{\partial \tilde{n}_H(t)}{\partial \tilde{n}(t)} = -\frac{1}{a_3} < 0\,; \quad \frac{\partial \tilde{n}_L(t)}{\partial \tilde{n}(t)} = \frac{a_1}{a_3} > 0\,; \quad \frac{\partial \tilde{n}_L(t)}{\partial \tilde{n}_H(t)} = -a_1 < 0\,;$$

$$\frac{\partial \tilde{n}_W(t)}{\partial \tilde{n}(t)} = -\frac{a_2}{a_3} < 0\,; \quad \frac{\partial \tilde{q}(t)}{\partial \tilde{n}(t)} = -\frac{a_4}{a_3} < 0\,.$$

Proposition 6. Along the adjustment path, (a) the numbers of high-efficiency and low-efficiency firms are, respectively, negatively and positively related to the total number of firms; (b) the rates of entry and exit are negatively related to the total number of firms.

We have analyzed thus far the dynamics of the system off the steady state, after it is initially displaced by some external shock. We now analyze how demand and cost shocks affect entry and exit at the initial period. Substituting $t = 0$, (38b) and (40) yield

$$n_W(0) = n_W^* + a_2(n_H^0 - n_H^*) \tag{43}$$

$$q(0) = n_W^* + a_4(n_H^0 - n_H^*) \,. \tag{44}$$

Note that $n_W^* = q^*$. Totally differentiating these equations with respect to ξ

$$\frac{dn_W(0)}{d\xi} = -a_2 \frac{dn_H^*}{d\xi} < 0 \,; \quad \frac{dq(0)}{d\xi} = -a_4 \frac{dn_H^*}{d\xi} < 0 \,,$$

where $dn_H^*/d\xi > 0$. Thus a permanent, unanticipated, outward demand shift implies less entry and exit in the short run. This is in contrast to its zero long-run impact on entry and exit.

The short-run effects of an unanticipated permanent fixed-cost shock are however ambiguous. Differentiating (43) and (44),

$$\frac{dn_W(0)}{dF} = \frac{dn_W^*}{dF} - a_2 \frac{dn_H^*}{dF} \gtrless 0; \quad \frac{dq(0)}{dF} = \frac{dn_W^*}{dF} - a_4 \frac{dn_H^*}{dF} \gtrless 0 \,.$$

The signs are ambiguous even when the full expressions for dn_W^*/dF and dn_H^*/dF are substituted. Hence a cost shock may have similar or opposite initial effects on entry and exit. However, they move in phase after the realization of the initial displacements.[7]

5. Concluding remarks

Industry evolution process is a highly complex phenomenon, influenced by growth and fluctuation of and interactions between heterogeneous firms, as well as entry and exit. Recent applied works in industrial organization have uncovered many interesting empirical regularities regarding the industry evolution process. It is perhaps fair to say that the developments of the industry evolution theory have lagged behind in explaining some of the stylized facts. This is probably not surprising, given that modelling any combination of firm heterogeneity, entry and exit, different sources of growth and fluctuations and the randomness inherent in them is intrinsically very hard. No single analytical model can hope to capture most of the

[7] The effects of temporary shocks can also be analyzed analogously. Compared with permanent shocks they are easier to trace insofar as steady-state values are unchanged. On the other hand, the task of definitizing (18)–(19) proceeds differently and hence the solution expressions are different from (37a) and (37b). The non-monotonicity property remains unchanged however.

important features of industry dynamics within a unified framework (unless one is willing to turn to a numerical approach).

The paper presents a simple industry dynamics model with entry adjustment costs: the sunk cost of entry increases with the number of entrants. These costs imply a non-instantaneous adjustment path to the steady state, so that it generates an analytical framework within which long-run and short-run dynamics can be characterized.

It is shown that the nature of the steady state depends crucially on the fixed entry costs as well as entry adjustment costs. The steady-state effects of demand and costs on entry–exit are different. The non-monotonicity of the value function emerges as a key property of the short-run dynamics, which implies non-monotonic paths of entry, exit, total number of firms and net entry. Entry and exit are positively related over time after the initial perturbation due to any external shock, *irrespective of the type of shock*. The positive correlation is due to two effects: a crowding-out effect and a future-value effect. The model thus provides an explanation for the observed non-monotonicity of total number of firms and net entry, and the positive correlation between entry and exit over time.

We should, however, caution that the model is not meant as representative of the industry evolution process. It abstracts from many other important features especially those which constitute sources of growth. Our motivation was to formulate a simple model that would be amenable to short-run as well as steady-state analysis while preserving some very basic features of industry dynamics such as firm heterogeneity and randomness in efficiency and that would provide an understanding of some of the empirical regularities. The model hopes to demonstrate that despite the complexities involved in formulating industry evolution models with heterogenous firms, it is possible to retain analytical tractability of long-run and short-run dynamics.

Appendix: Discussion of eigen roots

From Section 2,

$$h_n = 2\rho - 1 - (1-\rho)\left(\frac{c_H}{c_L} - 1\right)$$

$$g_v = -\frac{c_L}{\beta[(1-\rho)c_H - \rho c_L]}\,.$$

h_n and g_v as functions of ρ are plotted in Fig. A1.

Consider first the g_v function. At $\rho = 1$, $g_v = 1/\beta > 1$. It is discontinuous at $\rho_1 \equiv c_H/(c_H + c_L)$. Except at this point, it decreases with ρ. At $\rho = 0.5$,

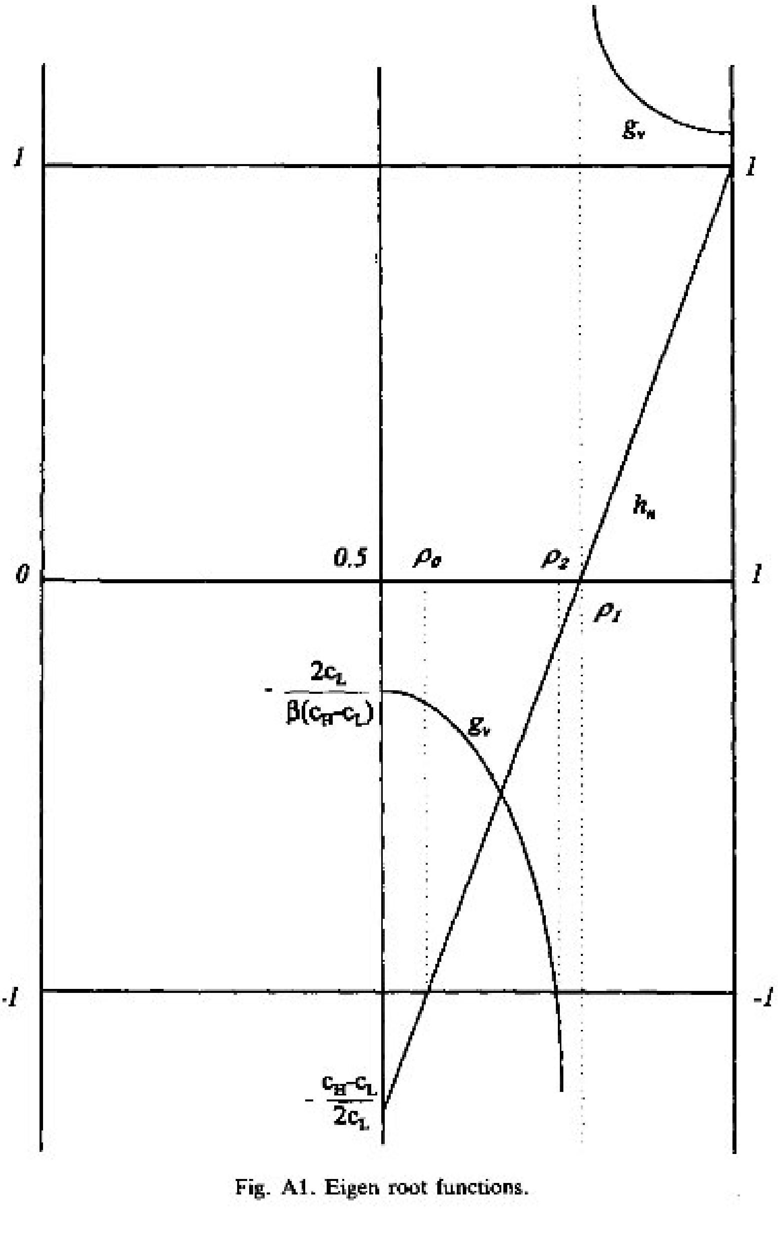

Fig. A1. Eigen root functions.

$$g_v|_{\rho=0.5} = -\frac{2c_L}{\beta(c_H - c_L)} < 0 .$$

In general, this term is greater or less than -1. But it increases with c_H/c_L. Thus if firm heterogeneity is sufficiently high, it is greater than -1. $g_v|_{\rho=0.5} > -1$ is equivalent to

$$\frac{c_H}{c_L} > \frac{2+\beta}{\beta} .$$

This motivated one of the two inequalities in Condition R. Given $c_H/c_L > (2+\beta)/\beta$, the g_v function behaves as depicted in Fig. A1.

Consider now the h_n function. It is linear and increasing in ρ. At $\rho = 1$, ρ_1 and $\rho_0 \equiv (c_H - c_L)/(c_H + c_L)$, respectively, h_n takes values 1, 0 and -1. At $\rho = 0.5$,

$$h_n|_{\rho=0.5} = -\frac{c_H - c_L}{2c_L} = \frac{1}{\beta g_v|_{\rho=0.5}} .$$

Hence, as long as $-1 < g_v|_{\rho=0.5} < 0$, $h_n|_{\rho=0.5} < -1$.

Defining another point $\rho_2 \equiv (c_H/c_L - 1/\beta)/(c_H/c_L + 1)$, and referring to Fig. A1, consider the following ranges of ρ. If $0.5 \leq \rho < \rho_0$, then $-1 < g_v < 0$ and $h_n < -1$. Thus a unique perfect foresight path exists with g_v as the root. If $\rho_0 < \rho < \rho_2$, then $-1 < h_n$, $g_v < 0$; a perfect foresight path exists but it is not unique. If $\rho_2 < \rho < 1$, then $|g_v| > 1$ and $|h_n| < 1$; a perfect foresight path exists with h_n as the stable root.

In summary, a unique perfect foresight path exists except for an intermediate range of ρ. However, as said before, the purpose of the model is to emphasize randomness in costs, not persistence or firm-specific effects, and moreover, in Section 3 it is shown that the greater the value of ρ, less likely is the existence of steady state with positive entry/exit and co-existence of H- and L-firms (see Fig. 1). Hence we are led to the values of ρ in the lower range in which a perfect foresight path exists, i.e. $0.5 \leq \rho < \rho_0$. $\rho < \rho_0$ is equivalent to

$$\frac{c_H}{c_L} > \frac{1+\rho}{1-\rho} ,$$

which is the other inequality in Condition R.

Thus under Condition R, a unique perfect foresight path exists with the (stable) root equal to g_v.

References

Baldwin, R.E., 1988, Hysteresis in import prices: the beachhead effect, American Economic Review 78, 773–785.

Baldwin, J.R. and P.K. Gorecki, 1991, Firm entry and Exit in the Canadian manufacturing sector, 1970–1982, Canadian Journal of Economics 24, 300–323.

Dixit, A., 1989, Hysteresis, import penetration, and exchange rate pass-through, Quarterly Journal of Economics CIV, 205–228.

Dunne, T. and M.J. Roberts, 1991, Variation in producer turnover across US manufacturing industries, in P.A. Geroski and J. Schwalbach (Editors), Entry and Market Contestability: An International Comparison (Blackwell, Oxford).

Dunne, T., M.J. Roberts and L. Samuelson, 1988, Patterns of firm entry and exit in U.S. manufacturing industries, Rand Journal of Economics 19, 495–515.

Ericson, R. and A. Pakes, 1989, An alternative theory of firm and industry dynamics, working paper, Columbia University and Yale University.

Geroski, P.A., 1991, Market Dynamics and Entry (Blackwell, Oxford).

Gort, M. and S. Klepper, 1982, Time paths in the diffusion of product innovations, Economic Journal 92, 630–653.

Hopenhayn, H., 1992, Entry, exit and firm dynamics in long run equilibrium, Econometrica 60, 1127–1150.

Jovanovic, B., 1982, Selection and the evolution of industry, Econometrica 50, 649–670.

Jovanovic, B. and S. Lach, 1989, Entry, exit, and diffusion with learning by doing, American Economic Review 79, 690–699.

Lambson, V.E., 1991, Industry evolution with sunk costs and uncertain market conditions, International Journal of Industrial Organization 9, 171–196.

Lambson, V.E., 1992, Competitive profits in the long run, Review of Economic Studies 59, 125–142.

Lippman, S.A. and R.P. Rumelt, 1982, Uncertain imitability: an analysis of interfirm differences in efficiency under competition, Bell Journal of Economics 13, 418–438.

Londregan, J., 1990, Entry and exit over the industry life cycle, Rand Journal of Economics 21, 446–458.

Schwalbach, J., 1991, Entry, concentration, and market contestability, in: P.A. Geroski and J. Schwalbach (Editors), Entry and Market Contestability: An International Comparison, (Blackwell, Oxford).