Review of Economic Studies (2006) **73**, 295–327 0034-6527/06/00110295$02.00

Firm Turnover in Imperfectly Competitive Markets[1]

MARCUS ASPLUND
London Business School and Columbia Business School

and

VOLKER NOCKE
University of Pennsylvania

First version received March 2004; *final version accepted June* 2005 (*Eds.*)

This paper is motivated by the empirical regularity that industries differ greatly in the level of firm turnover and that entry and exit rates are positively correlated across industries. Our objective is to investigate the effect of fixed costs and, in particular, market size on entry and exit rates and hence on the age distribution of firms.

We analyse a stochastic dynamic model of a monopolistically competitive industry. Each firm's efficiency is assumed to follow a Markov process. We show existence and uniqueness of a stationary equilibrium with simultaneous entry and exit: efficient firms survive, while inefficient ones leave the market and are replaced by new entrants. We perform comparative dynamics with respect to the level of fixed costs: entry costs are negatively related and fixed production costs positively related to entry and exit rates. A central empirical prediction of the model is that the level of firm turnover is increasing in market size. In larger markets, competition is endogenously more intense than in smaller markets, and so price-cost margins are smaller. This price competition effect implies that the marginal surviving firm has to be more efficient than in smaller markets. Hence, in larger markets, the expected lifespan of firms is shorter, and the age distribution of firms is first-order stochastically dominated by that in smaller markets.

In the empirical part, the prediction on market size and firm turnover is tested on an industry where firms compete in well-defined geographical markets of different sizes. Using data on hair salons in Sweden, we show that an increase in market size or fixed costs shifts the age distribution of firms towards younger firms, as predicted by the model.

1. INTRODUCTION

There is much simultaneous firm entry and exit going on at the industry level, and there is considerable variation in firm turnover across industries.[2] To explain these cross-industry differences in firm turnover is one of the important research agendas in the field of industrial economics. For example, Dunne, Roberts and Samuelson (1988) conclude their study as follows:

> The high correlation between entry and exit across industries indicates that industries differ substantially in their degree of firm turnover. One area for further study is then to identify the characteristics of industry technology and demand that give rise to across industry differences in turnover.

It is probably fair to say that the cross-industry differences in firm turnover are not yet very well understood. Although there are some notable exceptions (discussed below), there appear

1. A previous version of this paper, using a different data-set, was circulated and presented under the title "Imperfect Competition, Market Size, and Firm Turnover".

2. Caves (1998) provides a recent survey of the empirical literature on turnover and mobility of firms. See also Cabral (1997) and Sutton (1997*b*).

to be only a few theories that make empirically testable predictions regarding the determinants of firm turnover. This paper considers observable industry characteristics—fixed costs and, in particular, market size—and explores their effects on entry and exit rates in an imperfectly competitive industry. The same factors should also be expected to cause cross-industry variations in gross job reallocation.[3]

We analyse a stochastic dynamic model of an imperfectly competitive industry. Upon entry, a new firm gets an initial draw of its productive efficiency or the perceived quality of its product. Over time, a firm's efficiency is likely to change, as it is hit by idiosyncratic shocks. To illustrate, the perceived quality of a restaurant is partly given by the chef's skills. If the chef quits, the restaurant needs to hire a new chef, who may turn out to be better or worse than the previous one. In stationary equilibrium, firms follow a threshold exit policy: firms that are hit by a sufficiently bad shock optimally decide to leave the market and are replaced by new entrants. The focus of this paper is to relate the equilibrium rate of firm turnover to entry costs, fixed production (or opportunity) costs, and market size. Since the equilibrium distribution of firm efficiencies depends on industry characteristics, our model also provides testable predictions regarding the well-documented differences in productivity between and within industries.[4]

The central new prediction of the model is that the rate of firm turnover is increasing in market size. The result is based on the price effect of competition. In a free-entry equilibrium, a rise in market size causes the population of active firms to increase, which in turn leads to lower price-cost margins. Thus, there tend to be two opposing effects on firms' profits: larger sales (due to the increase in market size) and lower price-cost margins (due to the endogenous increase in the intensity of competition). In equilibrium, the net effect is positive for more efficient firms, but negative for less efficient firms, and so the marginal surviving firm has to be more efficient in larger markets. In larger markets, the expected lifespan of firms is, therefore, shorter, and the rate of firm turnover larger. It follows that the age distribution of firms in larger markets is shifted towards younger firms. Furthermore, economic integration of markets has selection effects that impact upon aggregate productivity: in larger markets, the endogenously more intense competition weeds out the less efficient firms and thus leads to a more efficient population of active firms.

Another prediction of our model is that firm turnover is lower in markets with higher entry costs. As entry costs increase, fewer firms find it profitable to enter the market, which reduces the competitive pressure, and allows less efficient firms to survive. An immediate implication is that regulations that increase entry costs (*e.g.* "red tape") not only reduce the number of entrants but also induce a less efficient population of firms. The effect of fixed costs is more subtle. An increase in fixed costs also makes entry less attractive, which corresponds to fewer firms in the market and less intense competition with higher equilibrium prices. There is, however, an opposing effect: each firm has to spend more on fixed costs, which reduces net profits. In equilibrium, the net effect is positive for more efficient firms, as they gain more from higher prices, and negative for less efficient firms. Consequently, the marginal surviving firm has to be more efficient in markets with higher fixed costs, and the expected lifespan of firms is shorter.

How can our predictions be tested empirically? The magnitude of the underlying fluctuations in the pattern of demand or technology is likely to vary greatly across industries. As pointed out by Sutton (1997*b*), this factor may be of primary importance, but it is very difficult to measure or to control for its impact empirically. This causes a serious problem for any empirical test of cross-industry predictions on firm turnover. Fortunately, by focusing on an industry with geographically independent ("local") markets, this problem can be largely circumvented. While

3. See Davis and Haltiwanger (1999) for a survey of the literature on gross job creation and destruction.
4. See Bartelsman and Doms (2000) for a review of the evidence.

local markets within the same industry differ in their size (and in firms' fixed costs), they share most of the unobservable and hard-to-control-for factors that would differ across industries. This is indeed the route taken in the empirical part of the paper, where we set out to test two predictions of our theory, using data on the age distribution of hair salons in Sweden: firms tend to be younger in (i) larger markets and (ii) markets with higher fixed costs. Using non-parametric tests of first-order stochastic dominance (FOSD) and regressions we find empirical support for both predictions.

The starting point of the recent literature on stochastic dynamic industry equilibria with heterogeneous firms is the seminal paper by Jovanovic (1982).[5] Jovanovic considers a perfectly competitive industry where firms have different but time-invariant efficiency levels. Firms only gradually learn their types over time by observing their "noisy" cost realizations. Firms that learn that they are efficient grow and survive, while firms that obtain consistently negative information decline and eventually leave the market. The model produces a rich array of empirical predictions on the relationship between firm growth and survival on the one hand and firm age and size on the other. However, all firms eventually learn their efficiency level, and so there is no firm turnover in the long run. Lambson (1991) considers another model with atomistic price takers. In his paper, there are no idiosyncratic shocks, but common shocks instead, to input prices (and demand). In equilibrium, firms may choose different technologies and hence be affected differently by the common shocks. The model predicts that the variability of firm values is negatively related to the level of sunk costs. Some empirical evidence for this prediction is given in Lambson and Jensen (1998) and Gschwandtner and Lambson (2002). Ericson and Pakes (1995) analyse a stochastic dynamic oligopoly model. There are two sources of uncertainty in their model: the outcomes of firms' investments in "quality" are stochastic and firms are subject to (negative) aggregate shocks. The equilibrium distribution of qualities at any time is itself stochastic. Few analytic restrictions can be placed on equilibrium outcomes. Instead, the authors have developed a simulation package; see Pakes and McGuire (1994).[6]

Hopenhayn (1992) is closely related to our model. The key difference between his model and ours is the assumed form of competition: Hopenhayn considers a perfectly competitive industry. The main prediction of his model is that firm turnover is negatively related to entry costs. Due to the absence of the price competition effect, however, market size has no effect on entry and exit rates. There are a few other papers that build on Hopenhayn's framework. Hopenhayn and Rogerson (1993) apply a general equilibrium version of the model to study the effect of changes in firing costs on total employment and welfare. Bergin and Bernhardt (1999) consider business cycle effects in a model of perfect competition. Both Das and Das (1997) and Melitz (2003) introduce monopolistic competition but assume a very special demand structure. Das and Das analyse the effect of entry adjustment costs on the convergence path to the stationary state. Melitz considers the impact of trade costs in a two-country version of Hopenhayn's model. In his model, the efficiency of incumbents does not vary over time, the death rate of incumbents is exogenously given, and the assumed demand structure does not allow for a price competition effect. Consequently, firm turnover is independent of market size.

The plan of the paper is as follows. In Section 2, we present the basic model. Then, in Section 3, we characterize (stationary) equilibrium and show existence and uniqueness. In Section 4, we investigate the comparative dynamics properties of the stationary equilibrium, which lie at

5. Note that we study the properties of a stationary industry equilibrium; we do not analyse the life cycle of an industry as in Klepper (1996). For more evidence on how entry and exit rates are related to the evolution of an industry, see also Carroll and Hannan (2000).

6. The passive learning model by Jovanovic (1982) differs from a number of other models (such as Hopenhayn, 1992; Ericson and Pakes, 1995; and our model) in that the stochastic process generating the size of a firm is non-ergodic. This is used by Pakes and Ericson (1998) to empirically distinguish between the two classes of models.

the heart of the paper. In the empirical part of the paper, Section 5, we investigate our prediction on market size and firm turnover using Swedish data on hair salons. Finally, we conclude in Section 6.

2. THE MODEL

We consider a stochastic dynamic model of an imperfectly competitive industry. There are a continuum of consumers and (potential) firms. Each firm produces a unique differentiated product and hence faces a downward-sloping demand curve. In a stochastic dynamic model, it is more convenient to work with a continuum of monopolistically competitive firms rather than with a finite number of oligopolistic players. First, if firms are atomless, we do not have to worry about integer constraints. In a free-entry equilibrium, the value of a new entrant is exactly equal to its outside option. Second, with a continuum of firms, idiosyncratic uncertainty washes out at the aggregate level. Hence, if uncertainty enters at the individual level only, all aggregate variables are deterministic. Third, the assumption of monopolistic competition greatly reduces the set of equilibria.

Firms differ in their "efficiency levels" or types that are subject to idiosyncratic shocks. Under our leading interpretation, the shocks directly affect firms' marginal costs. In this case, the marginal cost of each new entrant is independently drawn from the continuous cumulative distribution function $G(\cdot)$ with support $[0,1]$. An incumbent's marginal cost in period t, c_t, is given by

$$\begin{aligned} c_t &= c_{t-1} \quad \text{with probability } \alpha, \\ c_t &\sim G(\cdot) \quad \text{otherwise,} \end{aligned} \tag{1}$$

where $\alpha \in [0,1)$ measures the persistence of an incumbent's costs. The "shocks" to incumbents' efficiencies are assumed to be firm specific. There is, however, an alternative interpretation where firms differ in the perceived quality of their product; see Example 2 below. In this case, a firm's perceived quality (relative to marginal cost) is negatively related to firm type c. For brevity, we will refer to c as a firm's marginal cost in the remainder of the paper. The stochastic process (1) can be motivated by a simple economic model: the efficiency of a firm depends on the quality of its match with a manager that is distributed according to $G(\cdot)$. The manager stays with probability α, in which case the quality of the firm–manager match (and thus the firm's efficiency) remains unchanged. With probability $1-\alpha$, the manager leaves, and the firm has to get a new manager from the pool of managers. The advantage of this stochastic process is that it allows us to obtain tractable closed-form solutions. As a robustness check, we show in Asplund and Nocke (2003), that the main predictions of the paper do not rely on the particular Markov process assumed here; see our discussion at the end of Section 4.

Regarding sunk costs, we assume that a firm has to pay an irrecoverable entry fee $\epsilon > 0$ when it enters the market. Additionally, a firm faces a fixed production (or opportunity) cost of $\phi > 0$ per period.[7]

Time is discrete and is indexed by t. Firms have an infinite horizon and maximize the discounted sum of profits. The common discount factor is denoted by $\delta \in [0,1)$. In each period, the timing is as follows.

1. *Entry stage.* The potential entrants decide whether to enter the market or take up the outside option instead (the value of which is normalized to 0).

7. We could easily introduce a scrap value for exiting firms. However, any such scrap value would affect equilibrium in the same way as the fixed cost ϕ.

2. *Learning stage.* The entrants and the incumbents observe the realization of their current costs, c_t.
3. *Exit stage.* The new entrants and incumbents decide whether to leave the market forever (and take up the outside option).
4. *Output stage.* The active firms play some "market game", pay a fixed production cost ϕ, and receive profits.

Let us make two remarks on the sequence of moves. First, potential entrants decide whether or not to enter the market before knowing their current efficiency. This assumption is common to most dynamic industry models. It allows us to avoid assumptions about the size of the pool of potential entrants (other than that it is sufficiently large to ensure there is always a positive mass of firms that do not enter in equilibrium). This is of particular importance in our model as we are interested in the effects of market size. It seems plausible that the number of potential entrants is not independent of market size. Fortunately, with the assumed sequence of moves, we can remain agnostic about this relationship.[8] Second, new entrants are treated as incumbents at the exit stage, which takes place after firms have learnt their current types. This sequence of moves gives rise to a convenient and novel mathematical structure.

Formally, the model can be described as an *anonymous sequential game*; see Jovanovic and Rosenthal (1988). Let $\mathcal{M}$ denote the set of (Borel) measures on $[0,1]$ and $\mu \in \mathcal{M}$, the measure of firms' cost levels at the output stage. That is, for any interval $A \subset [0,1]$, $\mu(A)$ gives the mass of active firms with costs in A. We assume that all idiosyncratic uncertainty washes out at the aggregate level.[9] Hence, the evolution of the industry (from an arbitrary initial state μ_0) is deterministic.

To allow for a large class of models describing competition at the output stage, we do not model the "market game" explicitly. Instead of specifying firms' strategic variables (prices or quantities) and the details of the demand system, we make assumptions directly on firms' reduced-form profit function. Below, we show that these assumptions are satisfied by standard models of imperfect competition. Each firm's equilibrium profit is summarized by a reduced-form profit function. A firm's equilibrium profit depends on its own type, on market size, and on the endogenous distribution of active firms. The equilibrium profit (gross of fixed costs) of a type-c firm when the measure of firms is μ is denoted by

$$S\pi(c;\mu) \geq 0, \tag{2}$$

where S is a measure of market size (*e.g.* the mass of consumers in the market). By writing a firm's profit as in (2), we make a number of implicit assumptions. First, firms differ only in their (one-dimensional) types; they are symmetric in all other respects. Hence, if c denotes marginal costs, then the aggregate demand system is symmetric and competition non-localized. Second, an increase in market size means a replication of the population of consumers, leaving the distribution of preferences and income unchanged. Moreover, firms' marginal production costs are independent of output levels. This implies that market size enters (2) in a multiplicative way.[10] Of course, in equilibrium, the measure μ will depend on market size, but it is taken as given at the output stage.

8. In Nocke (2003), a different route is taken, which also avoids making arbitrary assumptions about the relationship between the size of the pool of potential entrants and market size. There, it is assumed that potential entrants, knowing their current type, self-select into markets of different size. Reassuringly, it is found that the central prediction of the present paper is insensitive to the assumptions on the entry process.

9. See Feldman and Gilles (1985) and Uhlig (1996) for a technical treatment of the problems associated with a law of large numbers for a continuum of random variables.

10. The assumption that market size enters multiplicatively into the gross profit function can easily be relaxed; see Nocke (2003).

To be more specific, let us consider the case where c denotes marginal costs. Denote by $SD(\cdot;\mu)$ and $P(\cdot/S;\mu)$ the demand and inverse demand functions, respectively, faced by an individual firm in equilibrium when the measure of active firms is given by μ. In this case, equilibrium gross profit can be written as

$$\begin{aligned} S\pi(c;\mu) &\equiv [p(c;\mu)-c]SD(p(c;\mu);\mu) \\ &= [P(q(c;\mu,S)/S;\mu)-c]q(c;\mu,S), \end{aligned} \tag{3}$$

where $p(c;\mu)$ and $q(c;\mu,S)$ are equilibrium price and output.

We will need to impose a few assumptions on the reduced-form profit function π. To motivate these, we first consider an example.

Example 1 (The linear demand model with a continuum of firms)**.** *There is a continuum* (*of mass* S) *of identical consumers with utility function*

$$U(\mathbf{x};H) = \int_0^n \left(x(i) - x^2(i) - 2\sigma \int_0^n x(j)x(i)dj \right) di + H,$$

where $x(i)$ is the consumption of variety $i \in [0,n]$, and H the consumption of the Hicksian composite commodity. The parameter $\sigma \in (0,1)$ measures the substitutability between different varieties. The linear–quadratic utility function gives rise to the well-known linear demand system.[11] *Each variety i is produced by a single firm at constant marginal cost $c(i) \in [0,1]$. The distribution of marginal costs in the population of active firms is summarized by the* (*Borel*) *measure μ. Firms compete in prices or, equivalently, quantities. In equilibrium, each firm faces the same residual demand curve, $SD(p;\mu) = S[\overline{c}(\mu) - p]/2$, where $\overline{c}(\mu)$ is implicitly defined by*

$$\overline{c}(\mu) = \frac{1+\sigma\int_0^{\overline{c}(\mu)} z\mu(dz)}{1+\sigma\mu([0,\overline{c}(\mu)])}. \tag{4}$$

The equilibrium gross profit of a firm with marginal cost c is then given by

$$S\pi(c;\mu) = \begin{cases} S[\overline{c}(\mu)-c]^2/8 & \text{if } c \le \overline{c}(\mu), \\ 0 & \text{otherwise.} \end{cases}$$

The gross profit function has thus the following properties. (1) *A firm's gross profit is strictly decreasing in the firm's marginal cost c, provided $c < \overline{c}(\mu)$. Firms with marginal costs larger than $\overline{c}(\mu)$ make zero gross profit.* (2) *Gross profits are proportional to market size S, holding fixed the population of firms.* (3) *Changes in the distribution of active firms affect profits through $\overline{c}(\mu)$, and the sign of the effect is the same for all firms. Hence, we can completely order firm distributions by their implied "intensity of competition"* (*which is inversely related to $\overline{c}(\mu)$*). (4) *"Adding" more active firms* (*with marginal costs below $\overline{c}(\mu)$*) *reduces firms' gross profits* (*by lowering $\overline{c}(\mu)$*). (5) *The gross profit function is continuous.* (6) *A change in the distribution of active firms* (*from μ to μ', say*) *which induces more intense competition* (i.e. $\overline{c}(\mu') < \overline{c}(\mu)$) *reduces the gross profit of more efficient firms by a larger total amount, but by a smaller fraction, than that of less efficient firms.*

The highlighted properties of the gross profit function are not specific to the linear demand model, but hold more generally in standard models of symmetric and non-localized competition.

11. This is the continuum version of the quadratic utility function which goes back to Bowley (1924). The associated demand system is widely used in oligopoly models; see Vives (1999).

In the Appendix, we show this for the case of the Cournot model (with homogeneous products and a finite number of firms which differ in their (constant) marginal costs of production).

Throughout the paper, we make the following assumptions on the reduced-form profit function.

(MON) *The reduced-form profit function $\pi(\cdot;\mu)$ is strictly decreasing in c on $[0,\overline{c}(\mu))$, and $\pi(c;\mu)=0$ for all $c\in[\overline{c}(\mu),1]$.*

This monotonicity assumption says that firms with lower marginal costs have higher profits. We allow, however, for the possibility that some inefficient firms (the types above $\overline{c}(\mu)$) may not sell their products even when offered at marginal cost and hence make zero gross profit.

Since the distribution of active firms is endogenous, we have to consider how changes in the distribution of firm types affect firms' profits. Intuitively, an "increase" in the population of firms should raise the intensity of price competition. Hence, we will say that a measure μ' is larger than another measure μ, $\mu'\succeq\mu$, if it induces more intense price competition and thus lower profits for all firms. Formally, we define a partial ordering, denoted by $\succeq$, on the set $\mathcal{M}$:

$$\mu'\succeq\mu \Leftrightarrow \quad \{\forall\, c\in[0,1],\ \ \pi(c;\mu')\leq\pi(c;\mu)\},$$

and

$$\mu'\succ\mu \Leftrightarrow \quad \{\mu'\succeq\mu,\ \text{and}\ \forall\, c\in[0,\overline{c}(\mu)),\ \ \pi(c;\mu')<\pi(c;\mu)\}.$$

We will say that measures μ' and μ are equivalent, $\mu'\sim\mu$, if they induce the same degree of price competition. The partial ordering implies $\overline{c}(\mu')\leq\overline{c}(\mu)$ for $\mu'\succeq\mu$.

(DOM) *If $\mu'([0,z])\geq\mu([0,z])$ for all $z\in(0,1]$, then $\mu'\succeq\mu$. If, in addition, the inequality is strict for some $z\in(0,\overline{c}(\mu))$, then $\mu'\succ\mu$.*

This assumption gives a sufficient condition for one distribution of firm types to induce more intense price competition than another. If the total mass of active firms remains unchanged, $\mu'([0,1])=\mu([0,1])$, (DOM) says that any shift in the population towards more efficient firms (in the sense of FOSD) makes competition more intense. More generally, (DOM) says that competition is more intense if the number of firms with marginal cost below z is larger, provided this holds for all z. We can remain completely agnostic about the effect on profits of increasing "average efficiency" of firms while reducing the mass of active firms. The role of assumption (DOM) in our model is to ensure that an increase in the number of entrants (holding the population of incumbents fixed) reduces the profit of any active firm and thus the value of an entrant.

(ORD) *The set $(\mathcal{M},\succeq)$ is completely ordered.*

Complete ordering of $(\mathcal{M},\succeq)$ is a common property of models of symmetric and non-localized competition. The assumption rules out that some firm makes larger profits when the distribution of active firms is given by μ rather than μ', while some other firm is better off under μ'. If firms differ only in their types and are symmetric otherwise, then all types should have the same ranking of distributions.

(CON) *The reduced-form profit function $\pi(c;\mu)$ is continuous.*[12]

For two of our main comparative dynamics results, we need to impose further structure on the reduced-form profit function. The following two conditions summarize properties of a large class of oligopoly models with heterogeneous firms.

12. Formally, we endow $\mathcal{M}$ with the topology of weak* convergence.

Assumption 1. *For $\mu' \succ \mu$, the profit difference $\pi(c;\mu) - \pi(c;\mu')$ is strictly decreasing in c on $[0, \overline{c}(\mu))$.*

Assumption 2. *For $\mu' \succ \mu$, the profit ratio $\pi(c;\mu')/\pi(c;\mu)$ is strictly decreasing in c on $[0, \overline{c}(\mu'))$.*

To explain these assumptions, consider an increase in the measure of active firms. By definition, this causes the gross profit of any firm to decrease, provided the firm makes a positive profit in the first place. Assumption 1 says that efficient firms suffer more than inefficient firms in terms of the absolute decrease in profit. Assumption 2 is the condition for our central result on the relationship between market size and market turbulence. It says that the percentage decrease in gross profit is larger for inefficient firms than for efficient ones. While both general and intuitive, these properties appear to have remained widely unnoticed in the literature.[13]

To further improve our understanding of Assumptions 1 and 2, the following proposition shows that they are equivalent to properties of equilibrium prices and quantities.

Proposition 1. *Suppose firm type c denotes marginal costs, and firms compete either in prices or in quantities. That is, equilibrium profit is given by* (3). *Assume also that the demand and inverse demand functions faced by an individual firm in equilibrium, $SD(\cdot;\mu)$ and $P(\cdot/S;\mu)$, are differentiable.*

1. *Assumption* 1 *holds if and only if equilibrium output $q(c;\mu,S) = SD(p(c;\mu);\mu)$ is decreasing in μ; that is, if and only if $q(c;\mu',S) < q(c;\mu,S)\ \forall c \in [0,\overline{c}(\mu)),\ \mu' \succ \mu$.*
2. *Assumption* 2 *holds if and only if the equilibrium price $p(c;\mu) = P(q(c;\mu,S)/S;\mu)$ is decreasing in μ; that is, if and only if $p(c;\mu') < p(c;\mu)\ \forall c \in [0,\overline{c}(\mu')),\ \mu' \succ \mu$.*

Proof. The assertions can be shown by taking the derivative of the profit difference and profit ratio with respect to c and applying the envelope theorem. ‖

By definition, an increase in the distribution of active firms causes the profits of firms to fall. For this to hold, the equilibrium price or quantity of any given firm must fall. If firm type c denotes marginal costs, Assumptions 1 and 2 say that both quantities and prices must fall in equilibrium. Proposition 1 is of independent interest. Since its proof does not make use of the fact that μ summarizes the distribution of active firms, it can be applied to any shift in consumers' tastes or incomes which reduces firms' profits.

In light of Proposition 1, Assumption 1 may be dubbed the *market share effect* of competition and Assumption 2 the *price effect* of competition. Assumptions 1 and 2 (as well as the other conditions we impose on the reduced-form profit function) are satisfied by a wide class of oligopoly models. Examples include the standard Cournot model with homogeneous products, and the linear demand model with a finite number of differentiated products and either price or quantity competition; in the Appendix this is shown for the Cournot model. As to models of monopolistic competition with a continuum of firms, the widely used Dixit–Stiglitz model satisfies most of our assumptions, including Assumption 1, but not Assumption 2. In fact, in the Dixit–Stiglitz model, there is no price competition effect: firms use a fixed mark-up pricing rule in which the mark-up is a function of some substitutability parameter in the utility function, but not of the population of active firms.

13. However, Boone (2000), in independent work, provides a few parametric examples to show that an increase in the toughness of competition results in an increase in the profit ratio between a more and a less efficient firm. This corresponds to Assumption 2.

While our discussion has focused on firm heterogeneity in marginal costs, there are other sources of firm heterogeneity that result in an isomorphic representation of the reduced-form profit function. To illustrate this, we give a version of the linear demand model where firms differ in their perceived qualities.

Example 2 (The linear demand model with perceived qualities). *This example is similar to Example* 1, *but all firms now have the same constant marginal cost* (*normalized to* 0). *The utility of the representative consumer is given by*

$$U(\mathbf{x};\mathbf{u};H)=\int_0^n\left(x(i)-\frac{x^2(i)}{u^2(i)}-2\sigma\int_0^n\frac{x(j)}{u(j)}\frac{x(i)}{u(i)}dj\right)di+H,$$

where $u(i)\in[1,\overline{u}]$ is the perceived quality of variety i, and the remaining notation is as before.[14] *Utility is strictly increasing in quality $u(i)$, provided that $x(i)>0$. In equilibrium, the gross profit of a firm offering quality u is given by $S(u-\underline{u}(\mu))^2/8$ if $u\geq\underline{u}(\mu)$, and equal to* 0 *otherwise, where the marginal type $\underline{u}(\mu)$ is implicitly defined by $\underline{u}(\mu)=\sigma\int_{\underline{u}(\mu)}^{\overline{u}}z\mu(dz)/[1+\sigma\mu([\underline{u}(\mu),\overline{u}])]$. Using the transformation $c=(\overline{u}-u)/(\overline{u}-1)$, the gross-profit function $S\pi(c;\mu)$ is of the same form as in Example* 1.

3. STATIONARY EQUILIBRIUM

We now turn to the equilibrium analysis. We are interested in analysing mature industries, and so we confine attention to a free-entry stationary equilibrium in which the distribution of firms is stationary. We characterize the stationary equilibrium, and show existence and uniqueness. We defer the analysis of the comparative dynamics properties, which lie at the heart of this paper, to the next section.

In a stationary equilibrium, the value of a type-c incumbent at the start of the exit stage, $V(c)$, can be written as

$$V(c)=\max\{0,\overline{V}(c)\},\tag{5}$$

where

$$\overline{V}(c)=[S\pi(c;\mu)-\phi]+\delta\left[\alpha V(c)+(1-\alpha)\int_0^1 V(z)G(dz)\right]$$

is the value conditional on staying in the market for one period and behaving optimally thereafter. (In other words, $\overline{V}(c)$ is the value of a type-c firm at the output stage.) The expression for the conditional value consists of two terms: the first term is the firm's current net profit $S\pi(c;\mu)-\phi$, the second captures the expected continuation value—with probability α, the firm will be of the same efficiency in the next period, while with the remaining probability it obtains a new draw from the distribution function $G(\cdot)$.

Let c^* be defined by

$$c^*\equiv\begin{cases}\sup\{c\in[0,1]\mid V(c)>0\} & \text{if } V(1)=0,\\ 1 & \text{if } V(1)>0.\end{cases}$$

Our assumptions on the profit function ensure that the value function $V(c)$ is continuous on $[0,1]$ and constant on $[\overline{c}(\mu),1]$. If $\pi(0;\mu)>0$, then standard arguments from dynamic programming imply that $V(c)$ is strictly decreasing on $[0,\min\{c^*,\overline{c}(\mu)\}]$. Hence, a firm's equilibrium exit strategy takes the form of a simple threshold rule, c^*, according to which the firm exits if and

14. This is the continuum version of the utility function in Sutton (1997*c*).

only if $c > c^*$. Let M denote the mass of entering firms in each period. In a stationary equilibrium,

$$\begin{gathered} V(c^*) \geq 0, \\ \text{with } V(c^*) = 0 \text{ if } M > 0. \end{gathered} \tag{6}$$

The value of an entrant at stage 1, V^{e}, is given by

$$V^{\mathrm{e}} = \int_0^1 V(c)G(dc) - \epsilon. \tag{7}$$

Free entry implies that

$$\begin{gathered} V^{\mathrm{e}} \leq 0, \\ \text{with } V^{\mathrm{e}} = 0 \text{ if } M > 0. \end{gathered} \tag{8}$$

Suppose the stationary equilibrium exhibits simultaneous entry and exit, that is, $M > 0$ and $c^* < 1$. In this case, $V^{\mathrm{e}} = 0$, and so $\int_0^1 V(c)G(dc) = \varepsilon$. This allows us to compute the conditional value function in closed form as

$$\overline{V}(c) = \begin{cases} \frac{S\pi(c;\mu) - \phi + \delta(1-\alpha)\epsilon}{1-\alpha\delta} & \text{if } c \leq c^*, \\ S\pi(c;\mu) - \phi + \delta(1-\alpha)\epsilon & \text{if } c \geq c^*. \end{cases} \tag{9}$$

In stationary equilibrium with simultaneous entry and exit, the free-entry condition $V^{\mathrm{e}} = 0$ becomes

$$\int_0^{c^*} [S\pi(c;\mu) - \phi + \delta(1-\alpha)\epsilon]G(dc) - (1-\alpha\delta)\epsilon = 0. \tag{E}$$

Similarly, the condition for optimal exit, $V(c^*) = 0$, becomes

$$S\pi(c^*;\mu) - \phi + \delta(1-\alpha)\epsilon = 0. \tag{X}$$

From exit condition (X), the current net profit of the marginal incumbent firm, $S\pi(c^*;\mu) - \phi$, is negative. This firm stays in the market only because there is a probability $1-\alpha$ that it will get a new draw from $G(\cdot)$, and the value of this option is $\delta(1-\alpha)\epsilon$. In a stationary free-entry equilibrium with simultaneous entry and exit, the value of a new draw must be equal to the entry cost ϵ. To simplify the exposition, we henceforth assume that $\phi \neq \delta(1-\alpha)\epsilon$. Then, a necessary condition for a stationary equilibrium with entry and exit is $\phi > \delta(1-\alpha)\epsilon$, which implies that the current gross profit of the marginal incumbent firm must be strictly positive, that is, $c^* < \overline{c}(\mu)$.

Let $\overline{V}^{\mathrm{e}}(x;\mu)$ denote the value of a new entrant who uses exit policy x in the period of entry and behaves optimally thereafter,

$$\overline{V}^{\mathrm{e}}(x;\mu) = \int_0^x \overline{V}(c)G(dc) - \epsilon. \tag{10}$$

In stationary equilibrium with simultaneous entry and exit, the free-entry condition (E) can then be rewritten as

$$\overline{V}^{\mathrm{e}}(c^*;\mu) = 0, \tag{E$'$}$$

and the condition for optimal exit as

$$\frac{\partial}{\partial x}\overline{V}^{\mathrm{e}}(c^*;\mu) = 0. \tag{X$'$}$$

The entry condition (E′) says that the value of an entrant who uses the equilibrium policy c^* in the first period, and behaves optimally thereafter, is equal to 0. The exit condition (X′) says that the derivative of this value function of an entrant with respect to the exit policy (and evaluated at the equilibrium exit policy c^*) is equal to 0. This close link between the entry and exit conditions

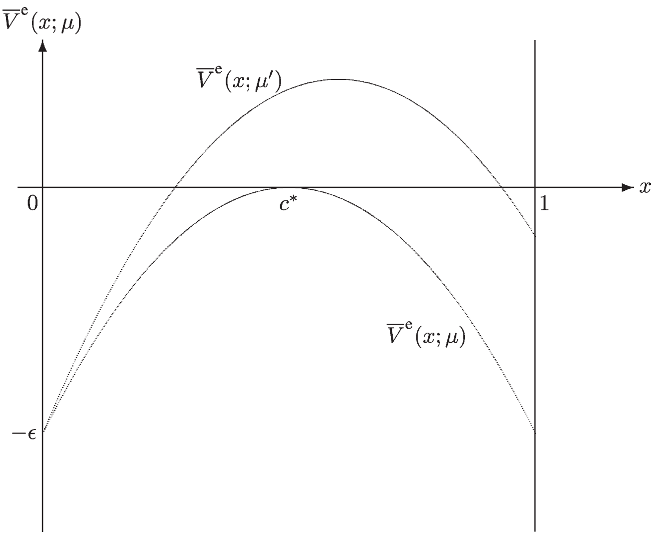

FIGURE 1

The effect of a decrease in the distribution of firms ($\mu' \prec \mu$) on the value of an entrant with exit policy x

is a consequence of the assumed sequence of moves. Since incumbents and new entrants face the same decision problem at the exit stage, the equilibrium exit policy of an incumbent, c^*, must maximize the value of an entrant. Hence, at the equilibrium exit policy, both the value of a new entrant and the derivative of the value function with respect to the exit policy have to be 0.

The properties of the value function $\overline{V}^{\mathrm{e}}(x;\mu)$ are straightforward to illustrate in Figure 1. First, $\overline{V}^{\mathrm{e}}$ is continuous in x and μ, and continuously differentiable in x. Second, and most importantly, $\overline{V}^{\mathrm{e}}(\cdot;\mu)$ is single-peaked on $[0,1]$. Clearly, if a new entrant immediately leaves the market no matter what its type, then its value is negative: $\overline{V}^{\mathrm{e}}(0;\mu) = -\epsilon < 0$. The entrant's value is increasing in its exit policy x as long as the value of an incumbent of type $c = x$ is positive; any further increase in the exit policy reduces the entrant's value. Hence, if μ gives the cost distribution in the stationary equilibrium, and c^* the equilibrium exit policy, then $\overline{V}^{\mathrm{e}}(\cdot;\mu)$ takes its unique maximum at $\widehat{x}(\mu) = c^*$. Finally, for all $x > 0$, the value $\overline{V}^{\mathrm{e}}(x;\cdot)$ is decreasing in the measure of active firms μ (provided $\overline{c}(\mu) > 0$).

The distribution of active firms is determined by firms' entry and exit decisions. Let $\mu[c^*, M]$ denote the invariant measure of firms' efficiencies at stage 4 if all firms follow exit policy $c^* \in (0,1)$, and the mass of entrants in each period is M. This measure is uniquely defined by

$$\mu[c^*, M]([0,z]) = \frac{M}{(1-\alpha)(1-G(c^*))} G(\min\{z, c^*\}), \qquad \forall\, z \in [0,1]. \tag{D}$$

The stationary distribution thus has the shape of distribution function $G(\cdot)$, is truncated at equilibrium exit policy c^*, and is scaled by the factor $M/[(1-\alpha)(1-G(c^*))]$. The stationary equilibrium with simultaneous entry and exit can now be defined as the triplet (μ, M, c^*) satisfying equations (E), (X), and (D).

In a stationary equilibrium *without* entry and exit, we have $M = 0$ and $c^* = 1$, and the stationary distribution is given by

$$\mu_\lambda([0,z]) = \lambda G(z), \qquad \forall\, z \in [0,1], \tag{11}$$

where $\lambda > 0$ is a scaling parameter. The stationary equilibrium without entry and exit is summarized by the triplet $(\mu_\lambda, 0, 1)$, satisfying (6), (8), and (11). Note that any stationary equilibrium distribution (with or without entry and exit) must be an element of $\mathcal{M}^* \equiv \{\mu \in \mathcal{M} \mid \forall z \in [0,1], \mu([0,z]) = kG(\min\{z,x\}),\ k > 0, x \in (0,1]\}$, that is, it has the shape of $G(\cdot)$, is scaled by some factor, and is truncated at some exit policy x.

Before we turn to the issues of existence and uniqueness of equilibrium, let us consider some features of the equilibrium that are consistent with stylized facts on industry dynamics. The probability of immediate exit of a new entrant is $1 - G(c^*)$, whereas an incumbent leaves the market with a smaller probability, namely, $(1-\alpha)(1-G(c^*))$. Empirical studies have indeed shown that new firms are the ones most likely to exit the market. Moreover, the model implies that new entrants are on average more efficient than exiting firms, but less efficient than surviving incumbents. This is again consistent with the empirical evidence. Finally, if firm size (*e.g.* measured by output) decreases with marginal cost c, the simple stochastic process given by (1) implies that firm growth is negatively related to firm size, as found by Evans (1987) and others.[15]

To simplify the proof of existence and uniqueness, we impose a technical condition on the reduced-form profit function π.

(FREE) *Fix any positive measure $\mu \in \mathcal{M}^*$, let $k > 0$ bca scaling parameter, and denote by μ^0 the null measure, that is, $\mu^0(A) \equiv 0$ for any Borel set A. Then,*

(i) $\lim_{k\to\infty} \pi(c; k\mu) = 0$ *for all $c > 0$, and*

(ii) $\int_0^1 S\pi(c;\mu^0)G(dc) > \phi + (1-\delta)\epsilon$.

Part (i) of the free-entry assumption (FREE) ensures that unlimited entry drives profits down to 0, while part (ii) implies that any stationary equilibrium distribution must be positive. Part (i) holds in standard models of competition, including the linear demand and Cournot models.[16] The same is true for part (ii), provided market size is sufficiently large relative to entry and fixed costs. We are now in a position to state our existence result.

Proposition 2. *There always exists a stationary equilibrium. Moreover, if a stationary equilibrium with simultaneous entry and exit exists, it is unique.*

Proof. See Appendix. ‖

There are two kinds of stationary equilibria: (i) with simultaneous entry and exit, and (ii) without simultaneous entry and exit. If the first kind exists, it is the unique equilibrium. If it does not, there exists an equilibrium in which a (positive) mass of firms is active, and no entry and exit take place. Since in this case, the entry and exit conditions become inequalities ($V^{\mathrm{e}} \leq 0$ and $V(1) \geq 0$, respectively), the mass of active firms in this equilibrium is not uniquely determined. In our proof of existence, we proceed in two steps. First, we neglect condition (D) and find the equilibrium exit policy c^* by varying the distribution of firms in $\mathcal{M}^*$ until (E) and (X) are satisfied. (If such a c^* does not exist in $(0,1)$, then the stationary equilibrium does not exhibit entry and exit.) Since $\overline{V}^{\mathrm{e}}(c;\cdot)$ is strictly decreasing in μ for $c \in (0,1]$, conditions (E) and (X) also pin down the equivalence class of the stationary distribution. Only then do we consider the

15. However, due to the simple Markov process (1), our model implies that the probability of exit of an incumbent is independent of the incumbent's age and size. This is not in line with much of the empirical evidence. However, in Asplund and Nocke (2003), we show that our results do not depend on the simple Markov process considered here, and so the model can be made consistent with the empirical evidence on hazard rates; see also our discussion at the end of Section 4.

16. To see this for the linear demand model, note that, from equation (4), $\lim_{k\to\infty} \overline{c}(k\mu) = 0$. For the Cournot model, see the Appendix.

stationary distribution generated by c^* and M, as given by (D). From assumptions (DOM) and (CON), and condition (D), it then follows that there exists a unique mass M of entrants such that the stationary distribution generated by M and the exit policy c^* is equivalent to the distribution determined in the first part of the proof. Note that our (novel) method of proof is applicable also to models of perfect competition where firms are price-takers.

It is straightforward to find conditions under which the unique stationary equilibrium involves simultaneous entry and exit.

Proposition 3. *If the entry cost ϵ is sufficiently small, there exists a unique stationary equilibrium with simultaneous entry and exit.*

Proof. See Appendix. ‖

For a stationary equilibrium to exhibit no entry and exit, the value of the least efficient incumbent (of type $c = 1$), conditional on staying in the market for another period, has to be larger than or equal to the value of an entrant. This incumbent is obviously less efficient than a potential entrant—but has already spent the entry cost. For small entry costs, the efficiency difference will outweigh the incumbent's sunk cost advantage, and so the entrant's value must be larger than the conditional value of the least efficient incumbent.

In the remainder of the paper, we will focus on the case of sufficiently small entry costs so that the stationary equilibrium exhibits firm turnover.

Remark 1. *While we are focusing here on the stationary equilibrium* (i.e. *on the steady state of the dynamic system*) *with simultaneous entry and exit, the* (*Markov*) *equilibrium* (μ_t, M_t, c_t^*), *starting from any initial distribution* μ_0, *can be defined in a very similar way. A natural question to ask then is whether the equilibrium will necessarily converge to the steady state. The answer is yes. Moreover, and perhaps more surprisingly, starting from any initial distribution* μ_0 *in period* 0, *there exists a finite date* T *such that, for all* $t \geq T$, μ_t *is equivalent* (*in terms of our ordering on measures*) *to the measure of active firms in the steady state. It follows that, for all* $t \geq T$, *the equilibrium exit policy* c_t^* *is the same as in the steady state. This means that the exit probabilities will coincide with those in the stationary equilibrium after finitely many periods. After date* T, *any further adjustment takes place through the measure of entrants,* M_t, *which will only asymptotically converge to its steady-state value.*

4. THE EFFECTS OF MARKET SIZE AND COSTS ON FIRM TURNOVER

In this section, we analyse the comparative dynamics properties of the stationary equilibrium. We begin by defining a measure of firm turnover and showing the close connection between firm turnover and the age distribution of firms. Then, we analyse the effect of changes in the level of the entry cost ϵ and the fixed cost ϕ on the equilibrium level of firm turnover. Next, we turn to the main concern of the paper, namely, the relationship between market size S on the one hand, and firm turnover and the age distribution of firms on the other.

4.1. *Measuring firm turnover*

A natural measure of (relative) firm turnover is the ratio between the mass of new entrants and the total mass of active firms in each period. This suggests defining the turnover rate θ as

$$\theta \equiv \frac{M}{\mu([0, 1])},$$

where the denominator is the mass of active firms at the output stage, and the numerator is the mass of new entrants.

Using (D), the turnover rate in the stationary equilibrium can be written as

$$\theta = (1-\alpha)\frac{1-G(c^*)}{G(c^*)}. \tag{12}$$

That is, given persistence α, there is a monotonically decreasing relationship between the equilibrium exit policy c^* and the turnover rate θ. The monotonicity of the relationship between c^* and θ not only holds for the simple Markov process (1) considered here, but for a large class of Markov processes, as shown in Asplund and Nocke (2003).

4.2. *Age distribution of firms*

In a stationary equilibrium, firms' exit policy shapes the age distribution of firms in a market. Let $\widehat{\theta}$ denote the (average) probability of exit of incumbents. In stationary equilibrium, $\widehat{\theta} = (1-\alpha)[1-G(c^*)] = G(c^*)\theta$.[17] We introduce the convention that the age of a new entrant (which is still active at the output stage) is one. Conditional on not leaving the market in the period of entry, the probability that a firm will survive until it reaches age a is then equal to $(1-\widehat{\theta})^{a-1}$. Consequently, the share of active firms whose age is less than or equal to a is given by

$$A(a|c^*) \equiv \frac{\sum_{t=0}^{a-1}(1-\widehat{\theta})^t}{\sum_{t=0}^{\infty}(1-\widehat{\theta})^t} = 1-(1-\widehat{\theta})^a,$$

which is strictly decreasing in c^*. Hence, a decrease in exit policy c^* shifts the age distribution towards younger firms in the sense of FOSD.[18]

4.3. *Entry costs and firm turnover*

Having defined a measure of firm turnover, we can now analyse the effects of changes in the parameters of the model on the turnover rate. Let us begin by looking at the effect of an increase in the level of entry costs.

Proposition 4. *An increase in the entry cost ϵ leads to an increase in exit policy c^*, and hence to a lower turnover rate θ and a shift in the age distribution of firms towards older firms. Furthermore, it causes the distribution of active firms, μ, and the mass of entrants per period, M, to decrease.*

Proof. See Appendix. ‖

The assertion of the proposition may be roughly explained as follows. Both the marginal incumbent (with cost level c^*) and the new entrant have a value of 0 in equilibrium. However, since incumbents have already sunk the entry cost, the average entrant has to be more efficient than the marginal incumbent. Clearly, this wedge in efficiency is increasing in the level of entry costs. That is, exit policy c^* increases with ϵ. This, in turn, implies that the hazard rate of incumbents is negatively related to the level of entry costs. A more formal explanation is the

17. The exit probability $\widehat{\theta}$ is another natural measure of firm turnover. The main results of the paper are not sensitive to the particular choice of turnover measure.

18. In Asplund and Nocke (2003), we show that this stochastic dominance result holds much more generally for a large class of Markov processes.

following. For any exit policy c^* and distribution μ, an increase in entry cost ϵ reduces the value of an entrant; that is, the curve $\overline{V}^{\mathrm{e}}$ in Figure 1 shifts downwards. For the entry condition to hold in the new equilibrium, the distribution of active firms must be smaller. This implies that the net profit $S\pi(c;\mu) - \phi$ of any type c goes up. This effect is reinforced by the fact that the option value of staying in the market rises with the level of entry cost. Hence, the value of the marginal incumbent in the initial equilibrium is now positive. Consequently, the marginal incumbent is less efficient in the new equilibrium. Hopenhayn (1992) derived the same prediction in a model with price-taking firms. A corollary of our result is that, in markets with higher entry costs, firms are on average less efficient, while the induced intensity of price competition is lower (in that the induced measure μ is smaller in the sense of our ordering on measures).

4.4. *Fixed costs and firm turnover*

Next, we analyse the effect of a change in ϕ, which may be interpreted as a fixed production cost or as an opportunity cost. The following proposition summarizes our results.

Proposition 5. *Suppose Assumption* 1 *holds. Then, an increase in the fixed* (*or opportunity*) *cost ϕ leads to a decrease in exit policy c^*, and hence to a higher turnover rate θ and a shift in the age distribution towards younger firms. Furthermore, it causes the distribution of active firms, μ, and the total mass of active firms, $\mu([0,1])$, to decrease.*

Proof. See Appendix. ‖

Holding fixed the distribution of active firms, an increase in the fixed cost ϕ shifts the curve $\overline{V}^{\mathrm{e}}$ downwards (Figure 1): for any exit policy, the value of an entrant decreases. Since the equilibrium value of an entrant is 0, this implies that, in the new equilibrium, the measure of active firms μ is smaller (in terms of our ordering on measures). Hence, there are two opposing effects on an incumbent's net profit $S\pi(c;\mu) - \phi$. On the one hand, the increase in fixed cost ϕ reduces the net profit of all types by the same amount. On the other hand, the endogenous decrease in μ reduces the intensity of price competition. All firms benefit from the higher equilibrium prices, but the market share effect (Assumption 1) implies that less efficient firms gain less in absolute terms (as they produce a smaller quantity). For the entry condition to continue to hold, the overall effect on profit must be positive for the most efficient firms, and negative for the least efficient active firms. In particular, the increase in fixed cost must decrease the profit of the marginal incumbent in the initial equilibrium. Hence, the exit policy c^* decreases with ϕ, which implies the predicted negative relationship between the fixed cost ϕ and turnover rate θ. Note that, in markets with higher fixed costs, the distribution of firm types is shifted towards more efficient firms, while the induced intensity of price competition is lower (in that μ is smaller in the sense of our ordering on measures). That is, if one were to correlate observable price-cost margins with the average efficiency level, one would find a *positive* cross-sectional correlation (assuming markets differed only in their level of fixed costs).

4.5. *Market size and firm turnover*

We now turn to our major concern, namely, the relationship between market size and firm turnover. The central prediction of this paper is summarized in the following proposition.

Proposition 6. *Suppose Assumption* 2 *holds. Then, an increase in market size S leads to a decrease in exit policy c^*, and hence to a rise in the turnover rate θ and a shift in the age* *distribution of firms towards younger firms. Furthermore, an increase in market size causes the distribution of active firms, μ, and the mass of entrants per period, M, to rise.*

Proof. See Appendix. ||

The result may be explained as follows. In a free-entry equilibrium, the measure of active firms, μ, is positively related to market size. Holding the distribution of active firms fixed, an increase in market size raises the value of firms for any exit policy. Graphically, this means that the curve $\overline{V}^{\mathrm{e}}$ in Figure 1 shifts upwards. Free entry then implies that the measure of firms has to increase with market size. This shifts the curve $\overline{V}^{\mathrm{e}}$ downwards. Hence, there are two opposing effects on a firm's gross profit $S\pi(c;\mu)$. On the one hand, the rise in market size S increases the profits of all firms proportionally. This can be thought of as an increase in output levels, holding prices fixed. On the other hand, as the distribution of firms increases, prices (and, hence, price-cost margins) fall. The price competition effect (Assumption 2) implies that the percentage decrease in profit from an increase in μ is greater the less efficient is the firm. Hence, if there is some type for which its value remains unchanged, then the value of all better types increases, and that of all worse types decreases. Since the equilibrium value of an entrant is 0, independently of market size, there must be some types in $[0, c^*]$, which are worse off in the larger market. In particular, the marginal incumbent in the smaller market would have a negative value in the larger market if it were forced to use the same exit policy as in the smaller market. This implies that the marginal surviving firm has to be more efficient in larger markets. That is, exit policy c^* is decreasing with market size. The hazard rate of the average incumbent firm is therefore higher in larger markets: firm turnover and market size are positively correlated. An immediate consequence is that the age distribution of firms in a smaller market first-order stochastically dominates that in a larger market. Moreover, in larger markets, the distribution of firm types is shifted towards more efficient firms, while the induced intensity of price competition is higher (in that μ is larger in the sense of our ordering on measures). That is, if one were to correlate price levels with the average efficiency level of active firms, one would find a *negative* cross-sectional correlation (assuming markets differed only in their size).

It is important to point out that our prediction on market size and firm turnover (Proposition 6) would not obtain in a model of perfect competition (as in Hopenhayn, 1992) or in a Dixit–Stiglitz type model of monopolistic competition (as in Melitz, 2003). The reason is that, in these models, the "price competition effect" (Sutton, 1997*a*) is absent, which implies that equilibrium price-cost margins are independent of market size.

The reasoning in the proof of Proposition 6 shows that efficient firms make higher profits in larger markets, and hence are more valuable. In contrast, less efficient firms are better off in smaller markets. That is, the distribution of profits and firm values is more "dispersed" in larger markets, while the distribution of efficiency levels is less "dispersed". The effect on profits is illustrated graphically in Figure 2.[19]

4.6. *Cost persistence and firm turnover*

What is the effect of changes in the persistence of costs, as measured by parameter α, on the turnover rate? In Asplund and Nocke (2000) we show that an increase in α has two opposing

19. Many oligopoly models imply that an increase in market size will lead to an increase in the number of firms and fiercer price competition. In a free-entry equilibrium, the lower margins will give rise to a concave relation between market size and the number of firms (see Bresnahan and Reiss, 1991). In Asplund and Nocke (2003), we show that in our model with differentiated products and an endogenous distribution of firm efficiencies, the concavity of the relationship is no longer guaranteed.

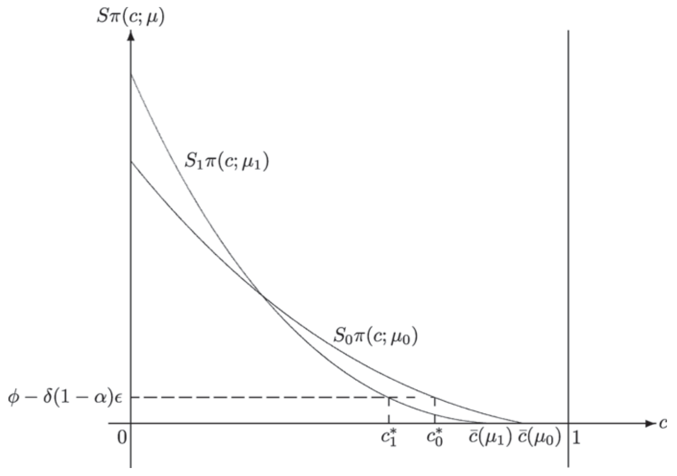

FIGURE 2

The effect of an increase of market size on gross profits: $S_1 > S_0$, and hence $\mu_1 \succ \mu_0$

effects on the turnover rate θ: the first term in (12), $(1-\alpha)$, clearly decreases with α, but the second term, $[1-G(c^*)]/G(c^*)$, is positively correlated with α. (The intuition for the latter observation is the following. The marginal incumbent makes a negative net profit, $S\pi(c^*;\mu)-\phi < 0$ and is in the market only because of the prospect of lower costs in the future ("option value"). An increase in the persistence of costs decreases the value of this option, and so the marginal incumbent has to be more efficient.) The overall effect on θ is ambiguous.[20]

For empirical work, the "non-neutrality" of α has important implications. In a cross-industry study of firm turnover, it may be difficult to control for differences in the stochastic process governing the evolution of firms' efficiencies or consumers' tastes (such as the persistence parameter α). To minimize measurement problems in this dimension, in the following section, we analyse turnover rates across independent local markets within the same industry.

4.7. *Robustness*

In Asplund and Nocke (2003), we explore the robustness of our results along several dimensions. First, we show that our results would remain unchanged if incumbents' exit decisions were to take place before they learn the realization of their current type. Second, we relax the assumption that the evolution of a firm's efficiency is governed by the simple Markov process (1). Specifically, for an incumbent whose current type is c, the cumulative distribution function of his type in the next period is given by $F(c'|c)$, where F is strictly increasing in its first argument and strictly decreasing in its second argument. In stationary equilibrium, larger firms have thus smaller hazard rates, which is strongly suggested by the empirical evidence.[21] Our existence and uniqueness results as well as the result on the effect of entry costs on firm turnover carry over to this more general setting. To prove our results on market size and fixed costs, we impose some additional structure on the Markov process; see Asplund and Nocke (2003) for details. A sufficient condition for an increase in the equilibrium exit policy c^* to first-order stochastically shift the age

20. However, an increase in α has an unambiguously negative effect on the exit probability $\widehat{\theta}$ (which does not keep track of those entrants that leave immediately after learning their current type).

21. By imposing further assumptions on F, one may also obtain that older firms are more likely to be larger and thus less likely to exit, which is again consistent with the empirical evidence.

distribution towards older firms is that the distribution functions F and G are such that the probability that a firm exits $t \geq 1$ periods after entry, conditional on having survived for $t-1$ periods, is decreasing in the exit policy c^*. Third, assuming that $\alpha = 0$, we analyse equilibrium when market size changes over time. Intuitively, one may expect entry rates to be higher in periods of growing market size and exit rates to be higher in periods of declining market size. This intuition is not correct, however, as our analysis reveals. In fact, we show that the predictions of our stationary model carry over to a large extent: exit rates are still increasing in the *level* of current market size; the same holds for entry rates, holding last period's market size fixed.

5. EMPIRICAL APPLICATION: HAIR SALONS IN SWEDEN

We now turn to the empirical application of our theory. As stressed in the Introduction, we have focused on the comparative dynamics properties of our model with respect to "observables" such as entry costs, fixed costs, and market size. There are, however, some measurement problems that make it difficult to directly test whether our model can explain inter-industry differences in turnover rates. One important variable in our model, namely, the magnitude of the underlying idiosyncratic shocks (as captured by the persistence parameter α in our benchmark formulation) is hard to measure or to control for, and is likely to vary across industries. There are also likely to be other factors that contribute to cross-industry differences in entry and exit patterns (*e.g.* financial constraints, regulations, industry life cycles) and that are left out of our model.

The empirical strategy chosen in this paper is instead to focus on a single industry with geographically independent ("local") markets and to use variations in local market conditions—in particular, market size and fixed costs—to explain differences in entry and exit rates. An identifying assumption of this approach, discussed further below, is that other factors remain constant across geographical markets. In the context of studying differences in turnover rates, we believe this assumption to be much more reasonable within an industry than across industries.

Our empirical strategy requires that the selected industry consists of many local markets that vary in size and in firms' fixed costs, and that in most of these local markets there is a reasonably large number of firms producing a well-defined but differentiated product. For this reason, we have chosen data on hair salons in Sweden. Competition among hair salons corresponds closely to the assumptions of our model. Even in small towns, there is typically a large number of hair salons to choose from. Products are clearly differentiated in terms of location and quality of service (which is closely tied to the skills and personalities of employees). The assumption of monopolistic competition seems therefore to hold good in this industry. Furthermore, the definition of the relevant industry is very straightforward: only hair salons offer hair cuts. Finally, casual observation suggests that there is a great deal of entry and exit of hair salons. An important source of idiosyncratic shocks is likely to be the turnover of employees.

The central prediction of our theory is that an increase in market size causes a rise in the turnover rate of firms, and hence a shift in the age distribution towards younger firms. We test this prediction as follows: we split our sample into small and large markets and then test if the age distribution of firms in the subsample of small markets first-order stochastically dominates that of firms in the subsample of large markets. Furthermore, we use information on land values to test our prediction that higher fixed costs result in higher rates of firm turnover, as evidenced by a shift in the age distribution towards younger firms. We begin by describing the data, and then turn to non-parametric tests of first-order stochastic dominance (FOSD). In addition, we use regressions to check the robustness of our results and examine some competing explanations.[22]

22. Our statistical tests of FOSD are more direct tests of the theoretical predictions than the regressions, in the sense that FOSD implies a difference in means but the reverse is not true. Regressions, however, allow us to control for other factors that may influence the age distribution.

TABLE 1

Descriptive statistics

	Mean	S.D.	Min	10th	25th	50th	75th	90th	Max	N
AGE	10·89	9·27	0·01	2	4	9	15	25	58	1030
MSIZE	72,014	94,541	355	2865	8271	27,763	82,788	277,522	285,981	1030
FIRMS	103	174	1	2	7	24	77	459	577	1030
RENT	3591	2649	209	1339	1942	2874	4370	6611	25,580	1003
RENTMEAN	4338	3790	559	1547	2112	3217	4810	7608	16,113	1030
POPDENSITY	755	1319	1	13	29	79	1041	4006	4006	1030
POPGROWTH	0·034	0·070	−0·148	−0·069	−0·010	0·052	0·098	0·101	0·294	1030
MIGRATION	0·927	0·262	0·287	0·648	0·716	0·926	1·069	1·175	2·116	1030
YOUNGPOP	0·277	0·039	0·198	0·232	0·244	0·269	0·301	0·349	0·380	1030

5.1. *Data*

The 2001 edition of the Swedish Yellow Pages lists 7243 hair salons, out of which we contacted 1100 by phone.[23] The sample contains information from interviews with 1030 of these; the remaining were either unwilling to participate or not possible to reach. The majority of hair salons in Sweden are small, single-establishment operations. Chains play only a minor role in the sector (our impression from studying the Yellow Pages and discussions with a number of hairdressers is that less than 5% of the salons are part of a chain), and we identify each establishment with a firm.

The measure of firm age is the number of years the salon has been established at the current location, AGE. The median age in the sample is 9 years (see Table 1) but the data show, as expected, a wide range of ages—with the 10th and 90th percentile at 2 and 25 years, respectively.[24]

Our theory is concerned with the effects of market size on firm turnover. Finding an appropriate measure of market size requires careful consideration. Since haircuts are not very costly and are purchased frequently, and most consumers have many nearby salons to choose from, consumers are unlikely to travel significant distances to purchase the service. This suggests to measure market size by the population living in an area and to assume that the population is solely served by the salons located within the same area.[25] The smallest areas that the Yellow Pages allow us to identify, and for which population figures are available, are the 8977 five-digit postal codes. At this level of market definition, however, we are unlikely to measure market size correctly since many consumers frequent salons outside the postal code where they reside. On the other hand, defining markets very broadly, for example, by municipalities (of which there are 289 in Sweden), runs into the opposite problem: a municipality is likely to contain several submarkets with little or no overlap. We have therefore decided to use an intermediate level of aggregation, namely, postal areas, to define market boundaries. Our measure of market size, MSIZE, is then the population living within a postal area. In Sweden, there are 1534 postal areas, ranging in size from small villages with less than 200 inhabitants up to the three largest postal areas Stockholm,

23. The selection was conducted as follows. Each hair salon was assigned a number, and in a first step we randomly selected 1000 of these. With such a non-stratified sample, most observations are from medium to large towns. To obtain greater representation from small markets, we randomly selected an additional 100 hair salons from the subsample of markets (postal areas) with less than 10 hair salons. Some hair salons may have chosen not to pay to appear in the Yellow Pages, but we believe that given the small cost (<SEK900≈ EUR100) these make up only a tiny fraction of all hair salons.

24. We also have information on whether the present owner had been previously established in the same area but at a different address; approximately 40% of the observations fall into this category. There is no significant correlation between having moved location on the one hand, and firm age, market size, and land values on the other.

25. The number of inhabitants in a market may not be a perfect measure of market size. Unfortunately, official income statistics are only broken down to the more aggregated municipality level. The same applies to demographic information on gender and age composition. Preferences, and thereby *per capita* demand, could potentially also vary across markets.

Gothenburg, and Malmo with more than 200,000 inhabitants each. In our sample, 368 postal areas are represented, with a median population of 7096 inhabitants. To measure the number of firms in a market, we use the number of hair salons listed in the Yellow Pages for a given postal area, FIRMS. The median is 11 hair salons in a postal area. Are postal areas a reasonable market definition for hair salons? To a first approximation, the number of consumers needed to make a hair salon viable should be roughly the same across markets, which would translate into a strong correlation between population and the number of firms. The raw correlation (for 1534 postal areas) between MSIZE and FIRMS is 0·92. In contrast, the corresponding correlation for postal codes is only 0·16. A simple example can explain this. Consider a typical medium-sized town, which is a single postal area with 20,000 inhabitants and 20 postal codes. The centre of the town has three postal codes, and the 17 other postal codes comprise residential suburbs. Although relatively few people live in the centre, many of the hair salons are located there and, as a consequence, the majority of the other postal codes do not have a single hair salon. This explains the low correlation between population and the number of firms in a postal code. Aggregating to the postal area level reflects our lack of information on exactly how market demand is geographically distributed. In larger towns (say, with a population above 100,000), there are often several centres, and so the postal area is too broad a market definition. To sum up, the population in a postal area is an imperfect measure of market size, but should be a sufficiently powerful measure to allow us to distinguish between small and large markets. Since measurement problems are likely to be most pronounced in large towns, we focus our analysis on small- and medium-sized markets.

Our paper is also concerned with the effects of fixed costs on firm turnover. For hair salons, the cost of floor space is an important source of fixed costs that varies considerably between, and even within, towns. The closest proxy for which public information exists is the average assessed value per square metre for commercial properties, RENT (collected to provide a basis for property taxation). These data are broken down by postal code. We base our tests on this low level of aggregation so as to be able to measure as exactly as possible the fixed costs a particular firm faces. With reference to the example above, rents vary not only between residential suburbs and the centre, but certainly also between the postal codes in the centre. For the 890 postal codes represented in the sample, the median is SEK 2720/m^2 and the 10th and 90th percentiles are 1240 and 6190, respectively. (On 1 January 2001, the inter-bank exchange rate of SEK to EUR was 0·11280.) For comparison, we also calculate a measure of the average fixed costs in a postal area, RENTMEAN, by using the total value of commercial properties in the constituent postal codes as weights of RENT.

A potential problem of identification is that rents tend to be increasing with town size. In the sample, the raw correlation between MSIZE and RENT is 0·62. Also, the cost of floor space is usually higher in the centre, partly since demand is greatest there. However, since rents are not only determined by the number of potential customers for hair salons, but by many other factors, it should in principle be possible to separate the effects of market size and fixed costs. To test this assertion we argue that, to a first approximation, the number of firms (i) rises less-than-proportionally with market size (see Bresnahan and Reiss, 1991)[26] and (ii) is decreasing in fixed costs (as predicted by Proposition 5). In Table 2, we report the results from a parsimonious Tobit specification with FIRMS as the dependent variable.[27] In (2:1), where the largest markets are excluded, the coefficient of MSIZE is positive and that of MSIZE^2 negative. The coefficient

26. While this is not guaranteed in our model with heterogeneous firms, numerical analysis (using the linear demand specification) indicates that this holds true, provided market size is not too small; see Asplund and Nocke (2003).

27. The specification follows from assuming that FIRMS/MSIZE decreases linearly in both MSIZE and RENTMEAN. The regression equation is obtained by multiplying both sides of $\frac{\text{FIRMS}}{\text{MSIZE}} = \beta_1 + \beta_2\,\text{MSIZE} + \beta_3\text{RENTMEAN} + \varepsilon$ with MSIZE and adding a constant. A Tobit model is used since many (mostly small) markets lack a hair salon.

TABLE 2

Tobit regressions

	FIRMS	FIRMS	FIRMS
	(2:1)	(2:2)	(2:3)
MSIZE	1·203***	1·183***	0·597***
	[0·054]	[0·038]	[0·018]
MSIZE ∗ MSIZE	−0·011***	0·0001	0·002***
	[0·002]	[0·0007]	[0·0001]
MSIZE ∗ RENTMEAN	−0·027***	−0·065***	0·044***
	[0·010]	[0·006]	[0·002]
Constant	−2·448***	−3·077***	−2·540***
	[0·146]	[0·179]	[0·200]
S.D.	2·644	4·198	5·726
	[0·066]	[0·101]	[0·137]
Sample	MSIZE < 25,000	MSIZE < 75,000	Full
N	1465	1523	1534
log L	2334·5	2890·3	3237·5
P-value	0·000	0·000	0·000

The number of firms related to market size and fixed costs. Standard errors in brackets.
*Significant at 10%.
*Significant at 5%.
***Significant at 1%.

of the interaction term MSIZE ∗ RENTMEAN is negative and significant, which implies that, conditional on market size, there are fewer firms where rents are high. These results are in line with our predictions. In (2:2), where we include all but the 11 largest markets, the coefficient of MSIZE ∗ RENTMEAN remains negative, but the one of MSIZE^2 becomes insignificant. Using the full sample, as in (2:3), produces markedly different coefficients. As noted above, problems of measuring market size in the large towns, and using a market average of rents, are the most likely explanations as to why including the largest markets may distort the estimates. It is also conceivable that we underestimate the size of the largest markets by ignoring the possibility that *per capita* demand is higher there (*e.g.* people in big towns go more often to the hairdresser). Overall, the results suggests that it is reasonable to focus on the small- and medium-sized markets so as to reduce measurement problems.

The control variables we use in the regressions are discussed in conjunction with the results in Table 4.

5.2. *Statistical tests of FOSD*

Before examining the statistical results, it is useful to first look at the raw data. Figure 3 shows the cumulative frequencies of AGE when the sample of firms in markets with less than 75,000 inhabitants is split into two equally large subsamples, “small” and “large” markets, on the basis of MSIZE. Over virtually the entire range of firm ages, the cumulative frequency in the set of large markets is above that in the set of small markets. This is in line with our prediction that firms in small markets tend to be older than those in larger markets. Figure 4 gives the corresponding frequency distribution of AGE when the sample is split into “low” and “high” rent markets, on the basis of RENT. Here, the youngest firms tend to be in the set of markets with high rents, as predicted by our theory. However, the gap between the two curves narrows with increasing age, and the two curves intersect to the left of AGE = 20; at higher age, the gap between the curves remains small.

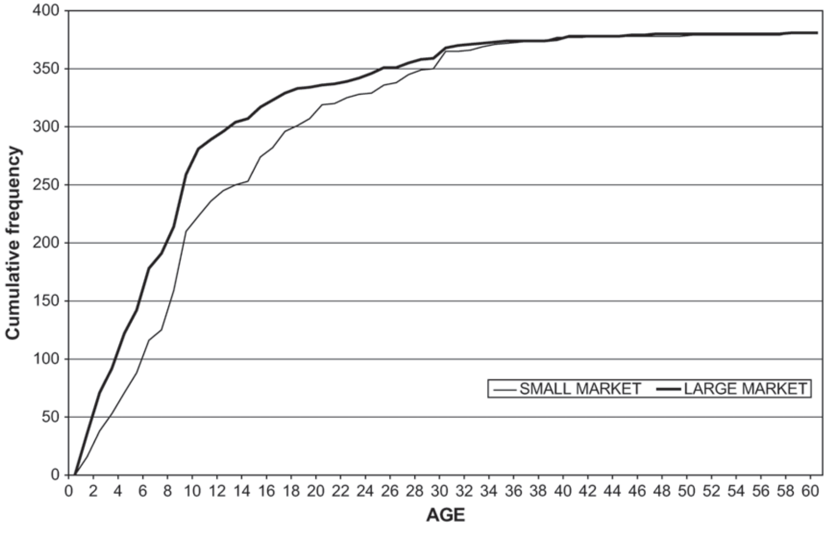

FIGURE 3

Cumulative age distributions in small and large markets

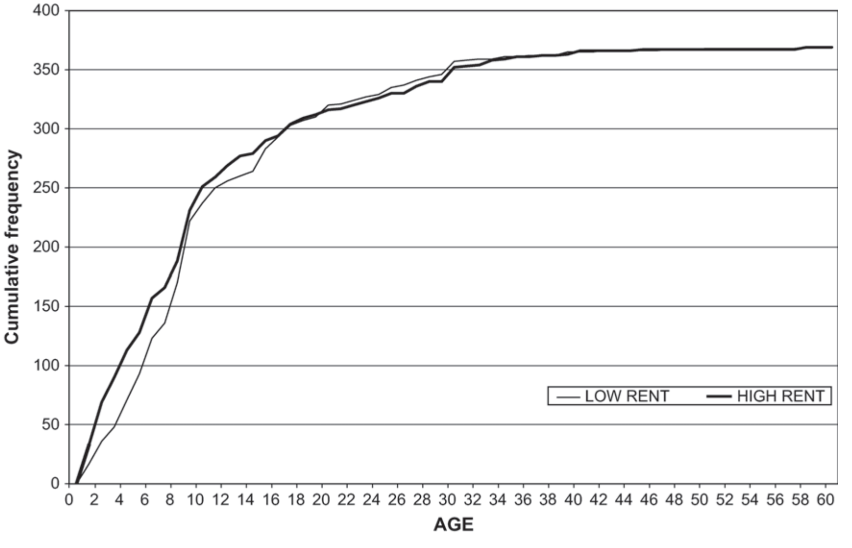

FIGURE 4

Cumulative age distributions in markets with low and high rents

To statistically test whether Figures 3 and 4 are consistent with FOSD, we use the non-parametric test proposed by Davidson and Duclos (2000). Let Y and Z be two random variables with cumulative distribution functions $F_Y(\cdot)$ and $F_Z(\cdot)$. Random variable Y first-order stochastically dominates Z, denoted $Y \succ_1 Z$, if

$$\begin{aligned} F_Y(a_i) &\leq F_Z(a_i) \quad \forall\, a_i, \text{ and} \\ F_Y(a_i) &\neq F_Z(a_i) \quad \text{for some } a_i. \end{aligned}$$

Suppose the data consist of N_Y and N_Z independent observations from $F_Y(\cdot)$ and $F_Z(\cdot)$, which form the empirical distributions, $\widehat{F}_Y(\cdot)$ and $\widehat{F}_Z(\cdot)$.[28] In the present context, $\widehat{F}_Y(\cdot)$ and $\widehat{F}_Z(\cdot)$ correspond to the proportions of firms that are less than or equal to a_i years old, in two different subsets of markets (small vs. large, and low rent vs. high rent markets). To test for FOSD, we compare the two distributions at a finite number of grid points a_i, $i = 1, \ldots, K$.

Davidson and Duclos show that, under the null hypothesis of equal distributions, $\widehat{F}_Y(a_i) - \widehat{F}_Z(a_i)$ is asymptotically normally distributed with variance

$$\begin{aligned}\widehat{V}(a_i) &= \widehat{V}_Y(a_i) + \widehat{V}_Z(a_i)\\ &= \frac{1}{N_Y}(\widehat{F}_Y(a_i) - (\widehat{F}_Y(a_i))^2) + \frac{1}{N_Z}(\widehat{F}_Z(a_i) - (\widehat{F}_Z(a_i))^2).\end{aligned}$$

The standardized test statistic is then given by

$$T(a_i) = \frac{\widehat{F}_Y(a_i) - \widehat{F}_Z(a_i)}{\sqrt[2]{\widehat{V}(a_i)}}.$$

Let $m_{\alpha,K,\infty}$ denote the critical value at the α per cent level of significance of the studentized maximum modulus test with K and infinite degrees of freedom, tabulated in Stoline and Ury (1979). We accept the hypothesis of $Y \succ_1 Z$ if

$$\begin{aligned}-T(a_i) &> m_{\alpha,K,\infty} \quad \text{for some } i, \text{ and}\\ T(a_i) &< m_{\alpha,K,\infty} \quad \forall\, i.\end{aligned}$$

If all $|T(\cdot)|$ are less than the critical value, which for $K = 10$ and $\alpha = 1, 5, 10, 20$ are 3·29, 2·80, 2·56, and 2·29, respectively, we accept the null hypothesis of equal distributions. In Table 3, all $T(\cdot) < 2{\cdot}29$, and so we refer to $|T| \equiv \max_i |T(a_i)|$.[29]

5.2.1. Test results. Table 3 shows the test statistic $T(a_i)$ at 10 grid points, which are chosen to divide the age distribution for the full sample into 11 intervals with an approximately equal number of observations. In the first two columns, we split the sample into two (equally large) subsamples of "small" and "large" markets. In the first column of Table 3, (3:1), we exclude firms in markets with more than 75,000 inhabitants. The maximum value of the test statistic is $|T| = 4{\cdot}85 > 3{\cdot}29 = m_{1,10,\infty}$. As predicted by our theory, the age distribution of firms in small markets first-order stochastically dominates that in large markets, and this is significant at the 1% level. Including all markets, as in (3:2), the test statistic becomes insignificant. One possible explanation for this is that the very largest towns contain a number of small submarkets, where competition is less intense (and hence firms are older) than in the city centre.

Since there may be measurement errors in MSIZE, which would imply a potential misclassification of markets by size, we report the range of the test statistic, denoted $T_{1/3}(a_i)$, when comparing the age distribution in the largest third of the markets with that in the smallest third.

28. There is a subtle issue whether the observations in our data are completely independently drawn. While we have randomly sampled about one in eight hair salons in Sweden, we drew each observation without returning it into the "urn" . Hence, sampling an observation from one market reduces the probability of drawing again from the same market. Moreover, the fact that one has drawn a firm of a certain age may change the expected age distribution of the firms from the same market that remain in the urn. We believe that this does not significantly reduce the power of our tests. First, we have sampled only a small fraction of firms, and (due to the large number of markets) most are from different markets. Indeed, we have rarely sampled more than one hair salon from a small market. Second, our simple Markov process (1) implies that, within the same market, a firm's age is independent of its efficiency, and hence independent of a rival firm's age.

29. Our results are almost identical when using the alternative test of FOSD proposed by Anderson (1996).

TABLE 3

Tests of first-order stochastic dominance

Split by	MSIZE (3:1)	MSIZE (3:2)	RENT (3:3)	RENT (3:4)
AGE	$T(a_i)$	$T(a_i)$	$T(a_i)$	$T(a_i)$
1	−2·889	−0·246	−2·352	−2·585
2	−3·441	−0·349	−3·506	**−3·327**
3	−3·545	0·015	**−4·008**	−2·975
5	−4·237	−1·068	−2·828	−3·309
7	**−4·855**	−2·196	−2·254	−2·485
9	−3·681	−1·570	−0·681	−1·131
12	−4·117	−1·946	−1·057	−0·924
15	−3·768	**−2·275**	−0·619	−1·157
20	−1·776	−1·010	0·427	0·263
35	−0·505	−0·867	0·000	−0·867
N	762	1030	738	1003
Range $T_{1/3}(a_i)$	[−3·74, −0·78]	[−3·26, −0·26]	[−3·48, 0·25]	[−3·16, −0·50]
Sample	MSIZE < 75,000	Full	MSIZE < 75,000	Full

Sample split by MSIZE and RENT. Critical values at 1%, 5%, 10%, and 20% levels: 3·29, 2·80, 2·56, and 2·29, respectively.

This split of the sample split should be less sensitive to potential misclassification. The maximum value of the test statistic in (3:2) is $|T_{1/3}| = 3{\cdot}26$; hence, the test statistic is significant even when including the largest markets.

The last two columns of Table 3 provide analogous test statistics for our variable RENT. In (3:3), we restrict attention to markets with less than 75,000 inhabitants, and find $|T| = 4{\cdot}01 > m_{1,10,\infty}$. As predicted by our theory, the age distribution of firms in low-rent markets first-order stochastically dominates that in high-rent markets, and this is significant at the 1% level.[30] The same is true in (3:4) for the full sample. The surprisingly significant test statistics in the full sample may be due to the fact that our variable RENT captures differences in rents across postal codes within the same postal area; in contrast, we assign the same value of MSIZE to all firms within the same postal area, even in large towns (which may explain the less significant results for MSIZE in the full sample).

It is encouraging that the results in Table 3 conform with the predictions of our theory. A potential caveat, however, is that MSIZE and RENT are highly collinear, and so it may not be too surprising to find similar results in the pairwise comparisons of small vs. large markets and low-rent vs. high-rent markets. To address this issue and to examine the robustness of our results, we now turn to a regression analysis.[31]

30. As seen in Figure 4, the two curves intersect to the left of AGE = 20, and one may ask whether our results are affected by the choice of grid points. The test statistic $T(a_i)$ assumes its maximum at AGE = 28, but $T(28) = 0{\cdot}874 < 2{\cdot}29 = m_{20,10,\infty}$. Thus, the violation of FOSD is statistically insignificant, and our conclusion is robust to the choice of grid points.

31. In Asplund and Nocke (2003), we address the collinearity issue by splitting the sample into four subsamples based on the sample medians of MSIZE and RENT: "small and low rent", "small and high rent", "large and low rent", and "large and high rent" markets. Our theory makes predictions for five bilateral comparisons of age distributions across these four subsamples. For instance, conditional on being in a small market, the age distribution of hair salons in a low rent area should first-order stochastically dominate that of salons in high rent areas. For each of the five bilateral comparisons, the test statistic $|T|$ is above the critical value at the 20% level, and most are above that of the 5% level.

TABLE 4

Least squares regressions

	ln(AGE) (4:1)	ln(AGE) (4:2)	ln(AGE) (4:3)
ln(MSIZE)	−0·126***	−0·125***	−0·125***
	[0·044]	[0·045]	[0·047]
ln(RENT)	−0·073	−0·103	−0·106
	[0·095]	[0·099]	[0·099]
ln(POPDENSITY)		−0·059*	−0·117**
		[0·033]	[0·047]
POPGROWTH		2·091**	1·788**
		[0·824]	[0·845]
MOBILITY			0·189
			[0·192]
YOUNGPOP			3·136
			[2·632]
Constant	2·342***	2·582***	1·831***
	[0·087]	[0·134]	[0·505]
Sample	MSIZE < 75,000	MSIZE < 75,000	MSIZE < 75,000
N	738	738	738
Adj. R^2	0·0217	0·0270	0·0289
Test 1	0·000	0·001	0·002
Test 2		0·041	0·018
Test 3			0·135

Age of firm related to market size and fixed costs. Robust standard errors in brackets. Observations clustered by market. Test 1 is the P-value of the restriction ln(MSIZE) = ln(RENT) = 0. Test 2 is the P-value of the restriction ln(POPDENSITY) = POPGROWTH = 0. Test 3 is the P-value of the restriction MOBILITY = YOUNGPOP = 0.
*Significant at 10%.
**Significant at 5%.
***Significant at 1%.

5.3. *Regression results*

In Table 4, we test our predictions by least squares regressions with ln(AGE) as dependent variable, focusing on markets with less than 75,000 inhabitants. In (4:1), we use only ln(MSIZE) and ln(RENT) as regressors. As expected, both coefficients are negative. However, only the coefficient on ln(MSIZE) is individually significant, which may be due to the collinearity of the two regressors. Since young and old hair salons coexist in almost all markets, only relatively little of the variation of ln(AGE) can be attributed to differences in market size and fixed costs across markets, which is reflected in the low explanatory power of the regressions.

5.3.1. Control variables. So far, we have implicitly assumed that markets differ only in terms of size and fixed costs. We acknowledge that conditions differ also in other dimensions that could influence the age distribution. We therefore now examine alternative explanations of our results.

Market size is not constant over time and the age distribution may reflect past changes in demand. For instance, consider two markets which are currently of the same size but which recently experienced different growth rates. One may expect firms to be, on average, younger in the market with the higher recent growth rate.[32] If large markets tend to be high growth markets, it

32. Assume that in one period, markets A and B both increase from size S_A and $S_B > S_A$, respectively, to size S and remain at this level forever after. Intuitively, since $c^* < c^*_B < c^*_A$, a greater fraction of firms in A will be non-viable

is possible that the tests of the market size effects reported above are only picking up this spurious correlation. Since no historical data are available on the population in postal areas, we use the growth rate of the municipality population from 1990 to 2000, POPGROWTH. Although there are fewer municipalities (289) than postal areas (1537), the municipality is still a quite narrowly defined geographical area, and the variable should be a reasonable proxy for demographic shifts that lead to changes in market size. In our sample, it is indeed the case that the large markets tend to have been in high growth areas—the correlation between MSIZE and POPGROWTH is 0·29.

Markets differ not only in population size but also in population density. Holding market size fixed, an increase in the population density is equivalent to a reduction in the geographical area, which implies that there is less room for spatial product differentiation: firms have to locate closer to each other, and hence competition (for a given population of firms) is more intense. Turnover rates may therefore not be independent of population density, holding market size fixed.[33] Information on population density at the postal area level does not exist, and so our variable POPDENSITY is measured at the municipality level.

To the extent that consumers visit hair salons outside the postal area in which they live, MSIZE does not fully reflect true market size. For a given value of MSIZE, this is more likely to be a problem in more densely populated areas (which, holding MSIZE fixed, encompass a smaller area). The variable POPDENSITY may therefore be picking up some of the effects of measuring market size imperfectly. In particular, if markets are more integrated in densely populated areas, then MSIZE would tend to underestimate the true market size.

Firms in some markets may experience more frequent shocks and will therefore tend to have younger firms. If the frequency of shocks is increasing in market size, then this could give rise to a picture like Figure 3. In our data, large markets tend to be in areas with relatively high turnover of people, and this could translate into greater propensity of employees to move or change employer as well as less stable customer relations. We attempt to control for this effect in our regressions by adding the variable MIGRATION, defined as the ratio of the sum of inward and outward migration over the period 1990–2000 to the municipal population in 2000. Another possibility, which is difficult to control for, is that the volatility in consumer preferences depends on market size, for example, due to the fraction of people swayed by fads and fashion being relatively high in cities. Possibly, this is related to there being more young people, whose preferences are more fickle, in areas with large markets. In order to capture this, we include the fraction of people aged 18–24 in the municipality, YOUNGPOP. Nevertheless, we believe that, especially outside the largest cities, which are excluded in our regressions, fads and fashion play a relatively minor role for the turnover of hair salons (as opposed to restaurants, for example).

5.3.2. Regression results with control variables. In the regression results reported in (4:2) and (4:3), the coefficients of our key variables, ln(MSIZE) and ln(RENT), are only affected to a limited extent by the inclusion of the additional control variables and their joint significance remains high. In (4:2), we include ln(POPDENSITY) and POPGROWTH as controls. The explanatory power in (4:2) is somewhat higher than in (4:1), and both ln(POPDENSITY) and POPGROWTH are individually significant; the two are jointly significant at the 5% level.[34] The

in the new larger market and will therefore exit. At the same time, the new larger market will support more firms and this will cause a greater increase in the inflow in the initially smaller market A. Both these effects, exit of old firms and entry of new firms, are stronger in market A, the market that experienced higher growth. A full examination of the effects of market size growth is outside the scope of this paper. In Asplund and Nocke (2003), we study the case of growing and declining market, assuming that there is no persistence in firms' efficiency levels ($\alpha = 0$).

33. In a two-period model, Syverson (2004*b*) shows that the productivity distribution of firms is related to the density of producers. In particular, the average productivity is higher when firms are more closely located.

34. We have also used 20-year population growth as a regressor and the coefficent remains positive but not statistically significant.

two additional controls in (4:3), MOBILITY and YOUNGPOP, are insignificant, both individually and jointly. The only noteworthy change compared to (4:2) is that the point estimate on ln(POPDENSITY) is lower. Overall, the regressions address the potential concern that the market size effect on the age distribution is driven by large markets being systematically different in other dimensions. Reassuringly, our results are *not* driven by large markets being those that have grown the fastest, have the most mobility, or have the youngest population.

While both market size and fixed costs shape the age distribution of firms as predicted by our theory, the effect of market size seems statistically stronger. One potential explanation for this is that our variable RENT is not only correlated with fixed costs, but also with other factors that impact the age distribution. In particular, it seems plausible that RENT is positively correlated with the level of sunk entry costs: upon entry, a salon is required to sign a lease contract for a minimum duration.[35] Furthermore, a new firm may need a certain time in the market to reach customers and to make an informed decision on exit. With the data at hand, we are not able to quantify the effects of RENT on fixed costs and entry costs.

Summing up, our study of the age distribution of hair salons in different local markets supports the predictions of our theory: firms operating in larger markets and those in locations where rents (being a proxy for fixed costs) are higher tend to be younger (in the sense of FOSD).

6. CONCLUSION

Many empirical studies in industrial organization and labour economics have shown that industries differ substantially in the level of firm turnover and gross job reallocation. These differences are stable over time and similar across countries. This suggests that there are some systematic factors that determine the magnitude of reallocation of inputs and outputs within industries. This paper is concerned with an examination of the role played by industry characteristics such as entry costs, fixed production costs, and market size on the turnover of firms.

To this end, we have analysed a stochastic dynamic model of an imperfectly competitive industry. Firms are heterogeneous and subject to idiosyncratic shocks to their "efficiencies". In our formulation, a firm's efficiency can be interpreted either as its productivity or as the perceived quality of its product.

Even in a stationary environment, the equilibrium exhibits simultaneous entry and exit: currently efficient firms survive while firms with sufficiently bad cost draws exit. Our analysis shows that the replacement of inefficient firms ("churning") is more rapid in markets with low entry costs and high fixed costs. The most important and novel prediction of our theory, however, is that the rate of firm turnover is related to the size of the market. As our analysis shows, the price effect of competition implies that large markets should have higher entry and exit rates than small markets.

In the empirical part of the paper, we test some of our model's predictions. The idea is to examine the age distribution of firms that compete in the same sector but in different geographical markets. To this end, we have collected data on hair salons in Sweden. We use the population in postal areas to capture differences in the size of markets and land values to proxy for differences in fixed costs (primarily rents). The empirical results are in line with the predictions of our model: hair salons tend to be older (in the sense of FOSD) in smaller markets and in markets with lower fixed costs.

35. Other costs for a hair salon do not appear to vary much across Sweden. According to the interest organization Sveriges Frisörföretagare, wages show very little dispersion across regions. Equipment and materials are typically bought from a few national distributors.

The empirical evidence from intra-industry data we provide suggests that cross-industry differences in market size and fixed costs may also help explain cross-industry differences in firm turnover. While it would be very interesting to test the predictions of our model using cross-industry data, any such study would need to address at least three difficult issues. First, our model shows that the turnover rate is related to the magnitude and frequency of underlying shocks to demand or costs, and it is hard to control for these in a cross-industry study. Second, it may be necessary to control for other factors that are left out of our model but that are likely to differ across industries (*e.g.* industry life-cycle effects). Third, it may not be straightforward to compare market size across industries. However, Sutton's (1991) work on the relationship between market size and concentration suggests that there is some hope of using market size in a cross-industry study.

This paper provides a number of testable implications in addition to those on firm turnover. It is empirically well-documented that firms within an industry display considerable heterogeneity in their efficiency levels. Our model makes a number of testable predictions on the distribution of productivities. In particular, the least efficient active firm is less efficient in smaller markets which implies a more dispersed distribution of efficiency levels. The only empirical study of plant level productivity across regional markets that we are aware of is Syverson's (2004*b*) examination of the ready-mixed concrete industry. His hypothesis is that as products in a market become closer substitutes, prices fall, which in turn makes it difficult for relatively inefficient firms to survive. He finds that demand density as well as the size of demand can help to explain moments of the efficiency distribution. The market size effect is statistically weaker, which he attributes to the inclusion of several control variables that are correlated with the level of demand. In our regressions, we found that both market size and population density were highly significant in explaining the age distribution of hair salons. Another prediction from our model is that there should be less dispersion of efficiencies when fixed costs are high and entry costs low. Syverson (2004*a*) explores variations in within-industry productivity for a sample of four-digit industries and finds some evidence that the dispersion of productivity is higher in industries where sunk costs are high and fixed costs are low.

Related to the distribution of efficiencies, it is often informally said that "more intense product-market competition fosters efficiency". This claim is true in our model in the following sense: in a cross-section of markets of different sizes, price levels are negatively correlated with the average efficiency of active firms. An increase in market size induces an increase in the number of active firms, holding the distribution of firms' efficiencies constant. But as the number of firms increases, prices fall, and so do margins. For some inefficient firms that were viable in a smaller market, the margins become so small that exit becomes the best alternative. Hence the large market is more competitive, in the sense that prices are lower, and the population of firms is on average more efficient.

The market size effect can also be related to the effects on industries that open up to trade: integrating previously independent markets will result in a larger market for firms' products. Melitz (2003) analyses the effect of trade costs on the distribution of efficiency levels in a two-country model. In his model, a move from infinite trade costs to 0 trade costs will have no effect on average efficiency. In contrast, a testable implication of our model is that full economic integration of two hitherto autarkic economies will change the average productivity as it results in a more efficient selection of firms. The same prediction would obtain in our model if foreign countries unilaterally open up (either partially or fully) to firms in the home country, increasing their effective size of the market.

Market integration will thus tend to have several positive effects on aggregate welfare: more efficient production, lower prices, and greater product variety. However, the higher turnover of firms implies a multiplication of entry costs. While Hopenhayn (1992) has shown that turnover levels in perfectly competitive industries are socially optimal, imperfectly competitive markets may exhibit excessive or insufficient turnover.

APPENDIX. PROOFS

A.1. *Properties of the profit function in a Cournot model*

We want to show that Assumptions 1 and 2 hold in a homogeneous goods Cournot model (with a finite number of firms), where firms differ in their (constant) marginal costs. (Of course, in the model, we assume that there is a continuum of firms producing differentiated products. The purpose of this section is to show that our assumptions on the reduced-form profit function reflect properties of standard oligopoly models. Our assumptions on profits are thus not specific to models with a continuum of firms, which we find reassuring. This is in contrast to the often-used Dixit–Stiglitz CES model, where the constant mark-up result holds only for a continuum of firms, but breaks down when the number of firms is finite.) Let $P(Q/S)$ denote inverse demand when aggregate output is Q and market size is S. We assume that the demand function is downward-sloping, that is, $P'(\cdot) < 0$. In equilibrium, aggregate output Q will be some (possibly complicated) function of the vector of firms' marginal costs, that is, $Q = Sf(\mathbf{c})$, where $\mathbf{c}$ is the vector of firms' marginal costs. It is, therefore, convenient to consider changes in aggregate output that reflect changes in the underlying distribution of firms' efficiencies. Conditional on aggregate output Q, the equilibrium output of a firm with constant marginal cost c is denoted by $q(c; Q, S)$. (The function $q(c; \cdot, S)$ is sometimes called the backward-reaction function.) The first-order condition for profit maximization is given by

$$P(Q/S) - c + \frac{q(c; Q, S)}{S} P'(Q/S) = 0, \tag{A.1}$$

which implies

$$q(c; Q, S) = -S\left(\frac{P(Q/S) - c}{P'(Q/S)}\right).$$

The associated second-order condition is given by

$$2P'(Q/S) + \frac{q(c; Q, S)}{S} P''(Q/S) < 0. \tag{A.2}$$

It is straightforward to show that (A.1) and (A.2) imply that a firm's equilibrium profit $S\pi(c; Q)$ is strictly decreasing in industry output Q. Any change in the underlying distribution of efficiencies, which reduces firms' profits, must induce an increase in industry output Q. Hence, an increase in Q is equivalent to an increase in the distribution of firms as defined in the main text. Moreover, the assumption of complete ordering of distributions (ORD) is satisfied. Conditional on industry output Q, the equilibrium profit of a type-c firm can be written as

$$S\pi(c; Q) = -S\frac{(P(Q/S) - c)^2}{P'(Q/S)}.$$

We now consider an increase in the distribution of active firms: suppose aggregate output increases from Q to Q', $Q' > Q$. Assumption 2 says that the profit ratio $\pi(c; Q')/\pi(c; Q)$ is decreasing in c for all c such that $q(c; Q, S) > 0$. It is immediate to see that this condition holds if and only if $P(Q'/S) < P(Q/S)$, which is clearly satisfied since the inverse demand function is downward-sloping. Assumption 1 requires that the profit difference $\pi(c; Q') - \pi(c; Q)$ is increasing in c for all c such that $q(c; Q, S) > 0$. Taking the derivative with respect to c, we obtain that Assumption 1 holds if and only if $q(c; Q', S) < q(c; Q, S)$, that is, a firm's equilibrium output is decreasing in the distribution of active firms. This inequality is satisfied if

$$P'(Q/S) + \frac{q(c; Q, S)}{S} P''(Q/S) < 0,$$

which is equivalent to the assumption that quantities are strategic substitutes (firms' reaction curves are downward-sloping). It is a fairly weak (and standard) assumption in Cournot models. Hence, in a Cournot model with homogeneous products and constant marginal costs, Assumptions 1 and 2 hold under fairly general conditions on demand. Further, it is easy to verify that assumptions (MON), (DOM), (ORD), and (CON) are satisfied as well. As regards (FREE), part (i) holds if $\lim_{Q\to\infty} [P(Q)]^2/P'(Q) = 0$, while part (ii) holds if $P(0) > 0$ and $[\phi + (1-\delta)\varepsilon]/S$ is sufficiently small.

Proof of Proposition 2. The proof proceeds in several steps.

Step one. Consider any positive measure μ' in $\mathcal{M}^*$ such that $\overline{V}^{\mathrm{e}}(1; \mu') = 0$. Conditions (CON) and (FREE) ensure that μ' exists. Since $\overline{V}^{\mathrm{e}}(\cdot; \mu')$ is single-peaked, there are two possibilities.

(i) $\overline{V}^{\mathrm{e}}(c; \mu') < 0\ \forall\, c \in [0, 1)$,
(ii) $\overline{V}^{\mathrm{e}}(c; \mu') > 0$ for some $c \in (0, 1)$.

Step two. In case (i), there does not exist a stationary equilibrium with simultaneous entry and exit. To see this, suppose otherwise that there exists a stationary equilibrium with simultaneous entry and exit. Denote the associated stationary distribution by μ'', and the exit policy by c''. Since $\overline{V}^{\mathrm{e}}(c;\mu)$ is decreasing in μ, condition (E) implies that $\mu'' \prec \mu'$. It follows that $\overline{V}^{\mathrm{e}}(1;\mu'') > \overline{V}^{\mathrm{e}}(c'';\mu'') = 0$. Since $\overline{V}^{\mathrm{e}}(\cdot;\mu'')$ is single-peaked, we thus have $\partial\overline{V}^{\mathrm{e}}(c'';\mu'')/\partial c > 0$, which contradicts condition (X). Although there does not exist a stationary equilibrium with simultaneous entry and exit, there does exist at least one without entry and exit. From (DOM) and (CON), it follows that there exists a positive number λ' such that $\mu_{\lambda'} \sim \mu'$, where $\mu_{\lambda'}$ is the stationary distribution defined by (11). (Indeed, (DOM) implies that for λ sufficiently large, $\mu_\lambda \succ \mu'$, and for λ sufficiently small, $\mu_\lambda \prec \mu'$. (CON) then implies that there exists a λ' such that $\mu_{\lambda'} \sim \mu'$.) It is easy to check that $\left(\mu_{\lambda'},0,1\right)$ satisfies conditions (6), (8), and (11). Note that there may exist a multiplicity of stationary equilibria. More precisely, there exists a non-empty interval of λ-values, $[\underline{\lambda},\overline{\lambda}]$, with $\overline{\lambda} \geq \underline{\lambda}$, such that $(\mu_\lambda,0,1)$, $\lambda \in [\underline{\lambda},\overline{\lambda}]$, forms a stationary equilibrium.

Step three. Consider case (ii), which can arise only if $\phi > \delta(1-\alpha)\epsilon$. We claim that, in this case, there exists a unique stationary equilibrium, which involves simultaneous entry and exit. Existence and uniqueness can be shown as follows. Starting from μ', we increase the measure of active firms: in Figure 1, this shifts the curve $\overline{V}^{\mathrm{e}}$ downwards. (CON), (FREE), and single-peakedness of $\overline{V}^{\mathrm{e}}(\cdot;\mu)$ imply that there exists a measure $\mu'' \succ \mu'$ such that $\overline{V}^{\mathrm{e}}(\cdot;\mu'')$ assumes a unique maximum at some $c^* < 1$ and $\overline{V}^{\mathrm{e}}(c^*;\mu'') = 0$. (It is easy to see that the exit policy c^* is unique.) From assumptions (DOM) and (CON), and condition (D), it follows that there exists a unique M such that $\mu[c^*,M] \sim \mu''$. The unique equilibrium distribution is then given by $\mu \equiv \mu[c^*,M]$.[36] ‖

Proof of Proposition 3. Abusing our previous notation, let us denote by $\overline{V}^{\mathrm{e}}(x;\mu;\varepsilon)$ the value of an entrant who uses exit policy x in the period of entry and behaves optimally thereafter, when the value of entry costs is given by ε. Note that $\overline{V}^{\mathrm{e}}(x;\mu;\varepsilon)$ is continuous (and decreasing) in ε. Assumptions (CON) and (FREE) ensure that there exists a positive measure $\widehat{\mu}$ in $\mathcal{M}^*$ such that $\overline{V}^{\mathrm{e}}(1;\widehat{\mu};0) = 0$. (MON) then implies that $S\pi(1;\widehat{\mu}) < \phi < S\pi(0;\widehat{\mu})$. It follows that there exists an exit policy $\widehat{c} < 1$, such that

$$\overline{V}^{\mathrm{e}}(\widehat{c};\widehat{\mu};0) > 0 = \overline{V}^{\mathrm{e}}(1;\widehat{\mu};0).$$

Hence, for ϵ sufficiently small, we have $\overline{V}^{\mathrm{e}}(\widehat{c};\widehat{\mu};\varepsilon) > 0 > \overline{V}^{\mathrm{e}}(1;\widehat{\mu};\varepsilon)$. Then, there exists a measure $\mu' \prec \widehat{\mu}$ such that

$$\overline{V}^{\mathrm{e}}(1;\mu';\varepsilon) = 0 < \overline{V}^{\mathrm{e}}(\widehat{c};\mu';\varepsilon).$$

That is, we are in case (ii) of the proof of Proposition 2. As we have already shown there, this implies that there exists a unique stationary equilibrium which involves simultaneous entry and exit. ‖

Proof of Proposition 4. Starting from a stationary equilibrium with simultaneous entry and exit, denoted by (μ_0,M_0,c_0^*), we consider an increase in the entry cost from ϵ_0 to $\epsilon_1 > \epsilon_0$. Let us assume that there is still a positive turnover rate in the new stationary equilibrium (μ_1,M_1,c_1^*), that is, $\theta_1 > 0$. (The proposition holds trivially if $\theta_1 = 0$.) From (9) and (10), it is easy to see that

$$\overline{V}^{\mathrm{e}}(x;\mu_0;\epsilon_1) < \overline{V}^{\mathrm{e}}(x;\mu_0;\epsilon_0)\ \forall\, x \in [0,1],$$

abusing notation by inserting argument ϵ into the value function $\overline{V}^{\mathrm{e}}$, as defined by (10). Hence, for condition (E) to hold in the new stationary equilibrium, we must have $\mu_1 \prec \mu_0$. Since $\mu_1 \prec \mu_0$ and $\epsilon_1 > \epsilon_0$, we must have $c_1^* > c_0^*$ for condition (X) to hold:

$$\begin{aligned} S\pi(c_0^*;\mu_1) - \phi + \delta(1-\alpha)\epsilon_1 &> S\pi(c_0^*;\mu_0) - \phi + \delta(1-\alpha)\epsilon_0 \\ &= 0 \\ &= S\pi(c_1^*;\mu_1) - \phi + \delta(1-\alpha)\epsilon_1. \end{aligned}$$

From (12), we thus obtain $\theta_1 < \theta_0$. To see that $M_1 < M_0$, notice that $\mu_1 \prec \mu_0$ and $c_1^* > c_0^*$, in conjunction with (DOM), imply that $\mu_1([0,c_0^*]) < \mu_0([0,c_0^*])$. Using (D), the result follows. ‖

Proof of Proposition 5. Starting from a stationary equilibrium with simultaneous entry and exit, denoted by (μ_0,M_0,c_0^*), we consider an increase in ϕ from ϕ_0 to ϕ_1, $\phi_1 > \phi_0$. Assume that there is still a positive turnover rate in the new stationary equilibrium (μ_1,M_1,c_1^*), that is $\theta_1 > 0$. From entry condition (E), it follows that

$$\overline{V}^{\mathrm{e}}(c^*;\mu_0;\phi_1) < \overline{V}^{\mathrm{e}}(c^*;\mu_0;\phi_0)\ \forall\, c^* \in (0,1],$$

36. If $\phi = \delta(1-\alpha)\epsilon$ (which occurs with zero probability if the parameters are drawn from some continuous distribution), two cases can arise: either $\overline{c}(\mu') = 1$ or $\overline{c}(\mu') < 1$. The former case is given by (i), that is, there exists a stationary equilibrium, but it does not exhibit entry and exit. In the latter case, $\overline{V}^{\mathrm{e}}(c;\mu') < 0$ for all $c \in [0,\overline{c}(\mu'))$ and $\overline{V}^{\mathrm{e}}(c;\mu') = 0$ for all $c \in [\overline{c}(\mu'),1]$. In this case, there exists both a stationary equilibrium without entry and exit ($c^* = 1$) as well as a continuum of equilibria with simultaneous entry and exit, where $c^* \in [\overline{c}(\mu'),1)$ and $\mu[c^*,M] \sim \mu'$.

and $\overline{V}^{\mathrm{e}}(0;\mu_0;\phi_1)=\overline{V}^{\mathrm{e}}(0;\mu_0;\phi_0)$. This implies that we must have $\mu_1 \prec \mu_0$ for (E) to hold again in the new stationary equilibrium. We now claim that $c_1^* < c_0^*$. To see this, suppose otherwise that $c_1^* \geq c_0^*$. According to condition (X),

$$S\pi(c_1^*;\mu_1)-\phi_1+\delta(1-\alpha)\epsilon = 0 = S\pi(c_0^*;\mu_0)-\phi_0+\delta(1-\alpha)\epsilon,$$

which implies that $S\pi(c_0^*;\mu_1)-\phi_1 \geq S\pi(c_0^*;\mu_0)-\phi_0$. Assumption 1 then ensures that

$$S\pi(c;\mu_1)-\phi_1 > S\pi(c;\mu_0)-\phi_0 \ \forall\, c\in[0,c_0^*). \tag{A.3}$$

Thus, we obtain

$$\begin{aligned}\overline{V}^{\mathrm{e}}(c_1^*;\mu_1;\phi_1) &\geq \overline{V}^{\mathrm{e}}(c_0^*;\mu_1;\phi_1)\\ &> \overline{V}^{\mathrm{e}}(c_0^*;\mu_0;\phi_0)\\ &= 0,\end{aligned}$$

where the first inequality follows from the fact that $\overline{V}^{\mathrm{e}}(\cdot;\mu_1;\phi_1)$ assumes a maximum at c_1^*, and the second inequality from (A.3). Now, $\overline{V}^{\mathrm{e}}(c_1^*;\mu_1;\phi_1) > 0$ cannot hold as it is in contradiction with (E). That is, we must have $c_1^* < c_0^*$. Finally, notice that, from (12), the turnover rate decreases monotonically with c^*, holding α fixed. Let us now show that we must indeed have $\theta_1 > 0$ (as assumed above), given that $\theta_0 > 0$. Define μ_0' and μ_1' by $\overline{V}^{\mathrm{e}}(1;\mu_0';\phi_0)=0$ and $\overline{V}^{\mathrm{e}}(1;\mu_1';\phi_1)=0$, respectively. (In the proof of Proposition 2, we have already shown that such measures exist.) It is easy to see that $\phi_1 > \phi_0$ implies $\mu_1' \prec \mu_0'$. Since $\theta_0 > 0$ by assumption, we have $\overline{V}^{\mathrm{e}}(c^*;\mu_0';\phi_0)=0$ for some $c^*\in(0,1)$. From Assumption 1, $\mu_1' \prec \mu_0'$, and $\phi_1 > \phi_0$, we get $\overline{V}^{\mathrm{e}}(c^*;\mu_1';\phi_1) > \overline{V}^{\mathrm{e}}(c^*;\mu_0';\phi_0)$ for all $c^*\in(0,1)$. This concludes the proof of the assertion on turnover. The result on the total mass of active firms follows immediately from $\mu_1 \prec \mu_0$ and $c_1^* < c_0^*$ (and (DOM)). ‖

Proof of Proposition 6. Starting from a stationary equilibrium with simultaneous entry and exit, denoted by (μ_0, M_0, c_0^*), we consider an increase in the size of the market from S_0 to $S_1 > S_0$. Let us assume that there is still a positive turnover rate in the new stationary equilibrium (μ_1, M_1, c_1^*), that is, $\theta_1 > 0$. (It is straightforward to show that turnover must indeed be positive in the new equilibrium, given that $\theta_0 > 0$. The argument is similar to that in the proof of Proposition 5, replacing Assumption 1 by Assumption 2.) The proof proceeds in several steps. First, notice that

$$\overline{V}^{\mathrm{e}}(x;\mu_0;S_1) > \overline{V}^{\mathrm{e}}(x;\mu_0;S_0)\ \forall\, x\in(0,1],$$

and $\overline{V}^{\mathrm{e}}(0;\mu_0;S_1)=\overline{V}^{\mathrm{e}}(0;\mu_0;S_0)$. For entry condition (E) to hold in the new equilibrium, we thus need $\mu_1 \succ \mu_0$. Second, suppose there exists a $y\in(0,\overline{c}(\mu_1))$ such that $S_1\pi(y;\mu_1)=S_0\pi(y;\mu_0)$. Assumption 2 then implies that $S_1\pi(c;\mu_1) > S_0\pi(c;\mu_0)$ for all $c\in[0,y)$, and the reverse inequality for all $c\in(y,\overline{c}(\mu_1))$. Third, assume the assertion of the proposition does not hold, and so $c_1^* \geq c_0^*$. Then,

$$\begin{aligned}S_1\pi(c_1^*;\mu_1) &= S_0\pi(c_0^*;\mu_0)\\ &\geq S_0\pi(c_1^*;\mu_0),\end{aligned}$$

where the equality follows from condition (X). From Assumption 2 we then obtain

$$S_1\pi(c;\mu_1) > S_0\pi(c;\mu_0)\ \forall\, c\in[0,c_1^*). \tag{A.4}$$

Consequently,

$$\begin{aligned}\overline{V}^{\mathrm{e}}(c_1^*;\mu_1;S_1) &\geq \overline{V}^{\mathrm{e}}(c_0^*;\mu_1;S_1)\\ &> \overline{V}^{\mathrm{e}}(c_0^*;\mu_0;S_0)\\ &= 0,\end{aligned}$$

where the first inequality follows from the fact that $\overline{V}^{\mathrm{e}}(\cdot;\mu_1;S_1)$ is maximized at c_1^*, and the second inequality from (A.4). Thus, entry condition (E) cannot hold in the new equilibrium: a contradiction. Consequently, we must have $c_1^* < c_0^*$, and hence $\theta_1 > \theta_0$. Observe that y exists and is in $(0,c_0^*)$; otherwise (E) would be violated. Finally, let us consider the effect of the increase in market size on the mass of firms that enter each period. Since $\mu_1 \succ \mu_0$ and $c_1^* < c_0^*$, we obtain (using (DOM)) $\mu_1([0,c_1^*]) > \mu_0([0,c_1^*])$, and hence, using (D), $M_1 > M_0$. ‖

Acknowledgements. We would like to thank the Managing Editor (Bernard Salanié), two referees, Boyan Jovanovic, Paul Klemperer, Steve Matthews, Meg Meyer, Rob Porter, Patrick Rey, Kevin Roberts, Mark Roberts, Larry Samuelson, Margaret Slade, and John Sutton, as well as seminar participants at several universities and conferences for helpful comments and discussions. Nocke gratefully acknowledges financial support by the National Science Foundation (grant SES-0422778).

REFERENCES

ANDERSON, G. (1996), "Nonparametric Tests of Stochastic Dominance in Income Distribution", *Econometrica*, **64** (5), 1183–1193.

ASPLUND, M. and NOCKE, V. (2000), "Imperfect Competition, Market Size and Firm Turnover" (CEPR Discussion Paper DP 2625, Centre for Economic Policy Research).

ASPLUND, M. and NOCKE, V. (2003), "Firm Turnover in Imperfectly Competitive Markets" (PIER Working Paper 03-010, University of Pennsylvania).

BARTELSMAN, E. J. and DOMS, M. (2000), "Understanding Productivity: Lessons from Longitudinal Microdata", *Journal of Economic Literature*, **38** (3), 569–595.

BERGIN, J. and BERNHARDT, D. (1999), "Industry Dynamics" (Mimeo, Queen's University and University of Illinois).

BOONE, J. (2000), "Competition" (Mimeo, Tilburg University).

BOWLEY, A. L. (1924) *The Mathematical Groundwork of Economics* (Oxford: Clarendon Press).

BRESNAHAN, T. F. and REISS, P. C. (1991), "Entry and Competition in Concentrated Markets", *Journal of Political Economy*, **99** (5), 977–1009.

CABRAL, L. M. B. (1997), "Competitive Industry Dynamics: A Selective Survey of Facts and Theories" (Mimeo, London Business School).

CARROLL, G. R. and HANNAN, M. T. (2000) *The Demography of Corporations and Industries* (Princeton, NJ: Princeton University Press).

CAVES, R. E. (1998), "Industrial Organization and New Findings on the Turnover and Mobility of Firms", *Journal of Economic Literature*, **36** (4), 1947–1982.

DAS, S. and DAS, S. P. (1997), "Dynamics of Entry and Exit of Firms in the Presence of Entry Adjustment Costs", *International Journal of Industrial Organization*, **15** (2), 217–241.

DAVIDSON, R. and DUCLOS, J.-Y. (2000), "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality", *Econometrica*, **68** (6), 1435–1464.

DAVIS, S. J. and HALTIWANGER, J. C. (1999), "Gross Job Flows", in O. Ashenfelter and D. Card (eds.) *Handbook of Labor Economics*, Vol. 3B (Amsterdam: North-Holland) 2711–2805.

DUNNE, T., ROBERTS, M. J. and SAMUELSON, L. (1988), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries", *Rand Journal of Economics*, **19** (4), 495–515.

ERICSON, R. and PAKES, A. (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work", *Review of Economic Studies*, **62** (1), 53–82.

EVANS, D. S. (1987), "Tests of Alternative Theories of Firm Growth". *Journal of Political Economy*, **95** (4), 657–674.

FELDMAN, M. and GILLES, C. (1985), "An Expository Note on Individual Risk Without Aggregate Uncertainty", *Journal of Economic Theory*, **35** (1), 26–32.

GSCHWANDTNER, A. and LAMBSON, V. E. (2002), "The Effects of Sunk Costs on Entry and Exit: Evidence from 36 Countries", *Economics Letters*, **77** (1), 109–115.

HOPENHAYN, H. A. (1992), "Entry, Exit, and Firm Dynamics in Long Run Equilibrium", *Econometrica*, **60** (5), 1127–1150.

HOPENHAYN, H. A. and ROGERSON, R. (1993), "Job Turnover and Policy Evaluation: A General Equilibrium Analysis", *Journal of Political Economy*, **101** (5), 915–938.

JOVANOVIC, B. (1982), "Selection and the Evolution of Industry", *Econometrica*, **50** (3), 649–670.

JOVANOVIC, B. and ROSENTHAL, R. W. (1988), "Anonymous Sequential Games", *Journal of Mathematical Economics*, **17** (1), 77–87.

KLEPPER, S. (1996), "Entry, Exit, Growth, and Innovation over the Product Life Cycle", *American Economic Review*, **86** (3), 562–583.

LAMBSON, V. E. (1991), "Industry Evolution with Sunk Costs and Uncertain Market Conditions", *International Journal of Industrial Organization*, **9** (2), 171–196.

LAMBSON, V. E. and JENSEN, F. E. (1998), "Sunk Costs and Firm Value Variability: Theory and Evidence", *American Economic Review*, **88** (1), 307–313.

MELITZ, M. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, **71** (6), 1695–1725.

NOCKE, V. (2003), "A Gap for Me: Entrepreneurs and Entry" (PIER Working Paper 03-019, University of Pennsylvania).

PAKES, A. and ERICSON, R. (1998), "Empirical Implications of Alternative Models of Firm Dynamics", *Journal of Economic Theory*, **79** (1), 1–45.

PAKES, A. and MCGUIRE, P. (1994), "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model", *Rand Journal of Economics*, **25** (4), 555–589.

STOLINE, M. R. and URY, H. A. (1979), "Tables of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons Among Means", *Technometrics*, **21** (1), 87–93.

SUTTON, J. (1991) *Sunk Costs and Market Structure* (Cambridge, MA: MIT Press).

SUTTON, J. (1997*a*), "Game-Theoretic Models of Market Structure", in D. M. Kreps and K. F. Wallis (eds.) *Advances in Economics and Econometrics: Theory and Applications. Seventh World Congress, Vol. 1, Econometric Society Monographs*, No. 26 (Cambridge: Cambridge University Press) 66–86.

SUTTON, J. (1997*b*), "Gibrat's Legacy", *Journal of Economic Literature*, **35** (1), 40–59.

SUTTON, J. (1997*c*), "One Smart Agent", *Rand Journal of Economics*, **28** (4), 605–628.

SYVERSON, C. (2004*a*), "Product Substitutability and Productivity Dispersion", *Review of Economics and Statistics*, **86** (2), 534–550.

SYVERSON, C. (2004*b*), "Market Structure and Productivity: A Concrete Example", *Journal of Political Economy*, **112** (6), 1181–1222.

UHLIG, H. (1996), "A Law of Large Numbers for Large Economies", *Economic Theory*, **8** (1), 41–50.

VIVES, X. (1999) *Oligopoly Pricing* (Cambridge, MA: MIT Press).