Typos

Industry Equilibrium Dynamics:

A General Competitive Theory.

Hugo A. Hopenhayn

March 1990

(Preliminary and incomplete version)

1. Introduction.

The study of industry dynamics has been an area of interest to both theorists and applied economists. Recent empirical studies indicate that a distinctive characteristic of industry evolution is the high degree of heterogeneity encountered: high variance of growth rates across firms, high dispersion in size and significant rates of turnover of firms.[1] Several dynamic models have been recently developed to account for this heterogeneity (see Jovanovic (1982), Lambson (1989), Ericson and Pakes (1989) and Hopenhayn (1989)). In spite of the similarities in the approach followed by the authors –all are equilibrium dynamic models– each of these models has its own specialized features and existence arguments.

The purpose of this paper is to provide a general theoretical framework for the study of competitive industry dynamics. As shown below, this framework accomodates a fairly large class of models, including those referenced above. A general set of assumptions are provided under which equilibria exist and the price dynamics is uniquely determined. As in Lucas and Prescott (1971), the equilibrium allocations are the solutions to an optimization problem, which under markovian assumptions is a dynamic program.

The theory presented depicts an industry composed by a large number of firms and potential entrants. Production possibilities of firms in each period, are affected by both aggregate and intrinsic sources of uncertainty. In any period, firms can exit the industry, while potential entrants may choose to enter. After entry and exit decisions have taken place, firms choose production plans for the period. These production plans affect their current profits and the conditional distribution for their state in the following period. Under

[1]See Dunne, Roberts and Samuelson (1987a,1987b), Evans(1987a,1987b) and Davis and Haltiwanger (1989).

the equilibrium price process all these decisions are optimal. The framework allows for aggregate and firm specific uncertainty, capital accumulation and the possibility of multi product firms. The technology sets of firms are not required to be convex, so U–shaped cost curves are admissible.

There are three distinctive features of the theory developed here: (i) Firms take the (correct) price process as given; (ii) there is a 'large number' of firms (more precisely a measure space of firms) which ensures the aggregate technology set is convex (though the individual technology sets of firms may not be so) and that all the aggregate uncertainty comes from the aggregate shocks; (iii) in any given period, all potential entrants are assumed to be ex–ante identical (though after entry they become differentiated by the realizations of the firm specific shocks).[2]

Of these, (i) and (ii) are more fundamental. For the finite dimensional case (e.g. finite number of periods or aggregate shocks), Novshek and Sonnenschein (1978) show that competitive industry equilibria where firms have infinitesimal size approximate Cournot competition when economies of scale are small relative to the size of the market. Small efficient scale also plays an important role in our setup. One may conjecture that the set of competitive equilibria described here approximate the Markov perfect equilibria of the corresponding dynamic game as economies of scale become small. But this is something we do not know.

[2]In fact, some degree of 'ex–ante heterogeneity' can be easily introduced, as mentioned in section 2.

From a methodological point of view, this paper extends Lucas and Prescott (1972) by allowing for heterogeneity, entry and exit and nonconvexities at the firm level. As in Lucas and Prescott, we show that industry equilibrium allocations maximize net expected discounted surplus and that allocations that maximize net expected discounted surplus and their corresponding prices are industry equilibria. There is an important difference, since there is no 'representative firm' in our framework, so the heterogeneity of the population of firms needs to be explicitly considered.

Section 2 introduces the elements that define an industry and describes the set of feasible allocations. Section 3 defines an industry equilibrium and lists the assumptions. Section 4 establishes the existence of equilibria and the uniqueness of equilibrium prices. Furthermore it develops the dynamic program that solves for the equilibrium allocations. Section 5 provides conditions under which the equilibrium is markov stationary. Section 6 studies the application of the theory to several models in the literature.

2. Definition of an Industry and Feasibility.

2.1. Industry

Before proceeding to a formal definition, we will (loosely) describe the basic elements involved. The industry considered is composed of a 'large' number of firms and potential entrants which at any point in time differ according to the value of an individual state, denoted by $s_t \in S$, where S is the state space of firms. This state may contain some exogenous (e.g. firm specific shocks) and some endogenous (e.g. capital) elements. Together with the realizations of an aggregate shock θ_t, the firm's specific state determines its production possibilities $Y(\theta_t,s_t)$, the set of input–output vectors that are feasible for the firm. In each period, the production plans of the incumbent firms, when aggregated, determine an aggregate input output vector, $y_t \in \mathbb{R}^n$ which is valued by the market according to a given function $p(\theta_t,y_t)$, where the j^{th} component $p_j(\theta_t,y_t)$ gives the market price for good $j=1,..,n$. For example, y_t could be a two element vector, the first one positive (output) and the second one negative (input), and in this case $p_1(\theta_t,\cdot)$ and $p_2(\theta_t,\cdot)$ are, respectively, the inverse demand function for the output and the inverse supply function for the input, which, as shown, may be affected by aggregate shocks.

The individual state of an incumbent firm for the following period is a function of its current state and input output choice, while possibly affected by the realizations of the aggregate shocks (θ^t), with conditional distribution given by $P(ds_{t+1}|\theta^t,s_t,y_t)$. Thus, before production decisions are made the aggregate shock (θ_t) and the firm specific state (s_t) determine the production possibilities of the firm in the current period and all relevant information for its future. Since it is the state of the firm and not its 'identity' which provides the relevant information[3], the situation of the industry at any point in time will be described by the distribution of firms across different states and, as a measure of the size of the industry (relative to the rest of the economy), the mass of firms in it. The industry starts with a measure μ_0 of firms, where $\mu_0(S)$ denotes the total mass of firms and $\mu_0(A)$ denotes the mass of firms with states $s_0 \in A$, where A is a measurable subset of S.

The possibility of entry (and exit) of firms to the industry, provides an extensive margin for changes in the aggregate input output vector of the industry. All potential entrants are assumed to be identical ex–ante, though after entry their states may differ. It is assumed that initial states of firms are independently drawn from a common distribution. This distribution may be affected by external and possibly time dependent factors (e.g. technological progress). This is modeled by making the initial distribution $\nu(ds|\theta_t)$ a function of the exogenous process $\{\theta_t\}$ faced by the industry. We are now ready for the formal definition.

An Industry is defined by the following set of elements:

$$\{p(\theta_t,y),\ \beta,\ Y(\theta_t,s),\ (S,\mathcal{S}),\ P(ds'\,|\,\theta_t,s,y),\ \nu(ds\,|\,\theta_t),\ \mu_0(ds)\}$$

where:

i. $\theta_t \in \Theta$ is a stochastic process common to all firms. θ^t denotes the history of the process up to time t and $\Psi(d\theta_{t+1}|\theta^t)$ its conditional distribution.

ii. $p:\Theta \times Y \rightarrow \mathbb{R}^k_+$ is an allocation pricing function, where $Y \subset \mathbb{R}^k$.

iii. $\beta \in [0,1)$ is a discount factor.

---

[3]"*In the limit the individual agent has so to speak lost his identity and it therefore seems artificial to keep the individualistic description of an economy in the limit*" (Hildenbrand, 1975).

iv. $Y(\theta_t,s) \subset Y$ is the technology correspondence which gives the technology set of a firm with own state 's' when the aggregate state is θ_t.

v. S is the space of states of the firms.

vi. $P(ds' | \theta_t,s,y)$ is the transition function for firm's states.

vii. $\nu(ds | \theta_t)$ gives the first period distribution of states for entrants.

viii. $\mu_0(ds)$ gives the initial distribution over states of firms in the industry.

As mentioned above, there are two possible sources of uncertainty: the aggregate process $\{\theta_t\}$ and the intrinsic uncertainty faced by each firm. Note that the aggregate stochastic process given in i. allows for non–stationarities, which can be useful to analyze the industry dynamics along its 'life cycle'.

Demand and supply conditions are given in ii. Following Arrow–Debreu conventions, inputs are denoted by negative numbers. Thus $p(\theta_t,y)$ gives the price vector of the goods produced and inputs used by the industry when the aggregate state is θ_t and the aggregate input output vector y.

Production possibilities for each firm are indexed by the aggregate state and by an individual state.[4] As mentioned above, the process for the state of the firms is given by an initial distribution $\nu(ds | \theta_t)$ from which the first period states for entrants are independently drawn and a transition function $P(ds';\theta_t,s,y)$, by which the states of firms are independently adjusted. Note that θ_t is an argument in the process followed by individual firms. So, for example, if there is an 'outside' innovation that makes all potential firms more productive, this could be captured in the above description by a change in the initial distribution faced by firms, e.g. a 'better' initial draw.

---

[4]The state space will typically be restricted to be a subset of a metric space. Hence the state of the firm could be a vector including exogenous components and capital stocks, a function, a measure or a combination of these.

The generality in the description of firms' states provides scope for many interesting special cases. In particular, 's' could be a vector with some purely stochastic components (that cannot be changed by the firm) while some components (e.g. capital, human capital) may be adjusted at a cost. It is also apparent that technologies with learning by doing –at the firm level– or 'time to build' are easily accommodated in this structure. For example, s could contain a vector of investments in progress, while some of the elements in y_t can be investment related inputs. The technology specified allows for 'fixed costs'; in that case $0 \notin Y(\theta_t, s)$. Entry costs can also be included by an appropriate restriction of the technology set for entrants, as detailed in section 3.2 .

A special case, widely used in the literature, is that of an industry with a homogeneous output, with price given by an aggregate demand function and production technology given by a cost function $c(q;\theta,s)$. By defining $Y(\theta,s) = \{(q,-c) | c \geq c(q;\theta,s)\}$ this problem translates into the above structure. It is worth emphasizing that the approach we follow has, as an advantage, that it generates explicit input factor demands.[5]

2.2 Feasibility.

As mentioned above, the distribution of firms' states and the 'size' of the industry, summarized in the measure μ_t, together with the aggregate process θ^t fully describe the state of the industry in period t. Hence input output choices of firms could be described by a function $y(\theta^t, s_t)$ with the obvious constraint $y(\theta^t, s_t) \in Y(\theta^t, s_t)$. However, this would restrict

[5]For an application of a model within this framework to the study of the effect of adjusment costs on labor turnover see Hopenhayn and Rogerson.

all firms with the same state to the same input output choice.[6] To allow 'different firms' to make 'different choices', instead of attaching to each state a unique input output vector, a distribution $\pi(dy;s)$ over feasible input output vectors can be specified. This distribution gives the fraction of firms in that state producing the various input output combinations.

An equivalent formulation which is analytically more convenient is to 'combine' the information on the input output choice of incumbent firms with the distribution of their respective states. This can be represented by a joint distribution over states and input output vectors (s_t,y_t) of firms, i.e. a measure $\mathscr{Y}_t$ on $S \times Y$. This measure satisfies $\mathscr{Y}(ds_t,dy_t)=\int \pi(dy;s)\mu_t(ds)$, where μ_t gives the distribution of incumbent firms' states and $\pi(ds';s)$ the conditional distribution of input output choices mentioned above. As noted, $\mathscr{Y}_t$ must be consistent with the distribution of incumbent firms' states μ_t, so its restriction to S (first marginal) must coincide with the measure μ_t. It must also specify allocations that are feasible for each state; more formally $\mathscr{Y}_t$ must have support on the graph of the correspondence $Y(\cdot)$, i.e. on pairs (s_t,y_t) such that $y_t \in Y(\theta^t,s_t)$.

The state of the industry is also affected by entry and exit decisions. Entry and exit occur at the beginning of each period, i.e. before production decisions are made. Since all entrants are ex–ante identical, the mass of entrants denoted by M_t summarizes entry decisions. Given that the distribution from which entrants independently draw their initial state is given by $\nu(ds|\theta^t)$, the contribution of entry in period t to the total distribution of firms is $M_t\nu(ds|\theta^t)$.

Exit occurs before firms observe there new state s_t but in knowledge of θ_t. At that point firms are distinguished not only by their state in the previous period s_{t-1} but also by

[6]This is a nontrivial restriction if the set of states were finite. In that case, if technology sets of firms are not convex, the aggregate production set will be nonconvex. In the existence proofs below aggregate convexity plays an important role.

y_{t-1}, their input output choice for that period; this is important since both elements affect the conditional distribution of the firm's state for the following period. Thus the firms that exit will be distinguished by (s_t, y_t) pairs with a measure of exits $\mathcal{E}_t$ on $S \times Y$ that satisfies the condition $\mathcal{E}_t \leq \mathcal{Y}_t$, or (loosely) exit of firms of type (s_t, y_t) cannot exceed the total available.

The allocations described above may of course depend on realizations of the process $\{\theta^t\}$. Thus the proper elements for the description of a feasible allocation are stochastic processes. The common information at time t, denoted by σ_t, is the minimal σ–algebra generated by the process $\{\theta^t\}$ in the usual manner.

A <u>feasible</u> <u>allocation</u> is a tuple of stochastic processes $\{\mathcal{Y}_t^*, M_t^*, \mathcal{E}_t^*, \mu_t^*\}_{t=0}^{\infty}$ (universally) measurable on $\{\Theta^t\}$, where $M_t^* \in \mathbb{R}_+$, $\mathcal{Y}_t^*$ is a measure on $S \times Y$, $\mathcal{E}_t^*$ is a measure on $S \times Y$, and μ_t^* is a measure on S, that satisfy the following conditions:

F1. $\mathcal{Y}_t^*$ has first marginal $\mu_t^* + M_t^* \nu$ and support on the graph of $Y(\theta^t, \cdot)$.

F2. $0 \leq \mathcal{E}_t^* \leq \mathcal{Y}_t^*$.

F3. $\{\mu_t^*\}_{t=0}^{\infty}$ is consistent with $\{\mathcal{E}_t^*, \mathcal{Y}_t^*, M_t^*\}_{t=0}^{\infty}$, i.e.
$\mu_{t+1}^*(ds') = \int P(ds' | \theta_{t+1}, s, y)[\mathcal{Y}_t^* - \mathcal{E}_t^*](dy, ds)$ and $\mu_0^* = \mu_0$.

We will use $\{y_t\}$ to denote the input output process that corresponds to a feasible allocation, which is derived from $\{\mathcal{Y}_t\}$ by the equation $y_t = \int y \mathcal{Y}(ds, dy)$. It is not hard to check that the set of feasible allocations is convex.

3. Industry Equilibrium.

3.1. Definition of Industry Equilibrium.

In any period, incumbent firms have two choices to make: whether to remain or leave the industry and in case they remain the choice of input output vector for the period. In making these decisions we assume firms have rational expectations and take as given the current prices and the distribution of future prices. Equilibrium prices for a given period depend on the realizations of the aggregate shock for that period and all previous ones. Potential entrants have the same information that incumbent firms have, except that their own states are not known prior to entry, though the distribution $\nu(ds;\theta_t)$ for their initial draw of s_t is anticipated. We will now proceed to the formal definitions.

A price process $p=\{p_t\}$ is a stochastic process on σ_t that takes values in $\mathbb{R}^n_+$. This process can be identified with a sequence of borel measurable functions from $p_t:\Theta^t \mapsto \mathbb{R}^n_+$. Given a price process $\{p_t\}$ the value of a firm when the aggregate state is θ^t and its own state s_t, is given by

$$v_t(\theta^t,s_t) = \sup_{T,y} E_t\{ \sum_{\tau=0}^{T} \beta^\tau p_{t+\tau} y_{t+\tau} \mid \theta^t,s_t\} \tag{3.1}$$

where both, the stopping time T and $y_t \in Y(\theta^t,s_t)$, are processes that are universally measurable on $\{(\Theta^t \times S)\}$. Note that $v(\theta^t,s_t)$ is the value of a firm in period t <u>after</u> observing the realization of its state and that the price process is implicit in this definition.[7]

---

[7]The requirement that the above be well defined is included in the definition of equilibrium.

A policy (τ,y) is <u>profit maximizing</u> for a firm with initial state s when the aggregate state is θ^t if it attains (3.1). Note that policies are stationary in the sense that τ and y do not depend on the whole history of states <u>of the firm</u> but only on its contemporaneous value. This restriction seems warranted since the process for the firm's 'idiosyncratic' shock is stationary and technology sets depend only on its current state. However, note that since the time variable 't' may be included in θ_t, a policy need not be time invariant.

An allocation $\{\mathcal{Y}_t^*, M_t^*, \mathcal{E}_t^*, \mu_t^*\}_{t=0}^{\infty}$ together with a price process $\{p_t^*\}$ is an <u>industry equilibrium</u> for industry $\mathcal{J}$ if:

(i) The allocation is feasible from μ_0

(ii) For all s in the support of μ_t and θ^t, (3.1) is well defined and $\mathcal{Y}(\cdot\,|\,\theta^t)$ has support on pairs (s,y) such that y is part of a profit maximizing policy under (θ^t,s)

(iii) a) For all firms in the support of $\mathcal{E}_t^*$, exit at t is part of a profit maximizing policy.

b) For all firms in the support of $\mathcal{Y}_t^* - \mathcal{E}_t^*$, remaining in the industry at time t is part of a profit maximizing policy.

(iv) $\int v(\theta^t,s)\nu(ds\,|\,\theta^t) \leq 0$ and if the inequality is strict $M_t^*(\theta^t)=0$.

(v) $p_t^* = p(\theta^t, y_t)$, where y_t is as defined in F1.

Conditions (ii) and (iii) state that firms maximize expected discounted profits, so their decisions are consistent with the objective (3.1); (iv) is the standard 'zero profit' condition for entrants: expected discounted profits of entrants cannot be strictly positive, and can only be negative when there is no entry; (v) is the standard market clearing condition.

3.2 Examples.

To better motivate the definitions we provide two examples. A detailed discussion of these and others is provided in section 6.

Lucas and Prescott (1971) analyze a dynamic stochastic competitive equilibrium for an industry with the following features: All firms have the same constant returns to scale technology with output constrained by capacity, i.e. if k_t is the firm's capacity at time t and q_t its output, then $q_t \leq k_t$. The cost of increasing capacity in one period from k_t to k_{t+1} is given by $g(k_t,k_{t+1}) = k_t h(k_{t+1}/k_t)$, where h is assumed convex, nonnegative and strictly increasing. Demand is given by a continuous demand function $D(\theta_t,q_t)$ decreasing in q_t, where θ_t follows a Markov process that takes values in $\mathbb{R}$.

Since there are no other factors that affect firms' production possibilities, the state of the firm is given by its stock of capital k_t. Entry is ruled out and since $0 \in Y(\theta,k_t)$ for all (θ,k_t), no exit of firms occur, so all changes in aggregate output and capital stocks are the result of changes in these variables by incumbent firms only. To model this in our framework, let $i_t=g(k_t,k_{t+1})$. Since g is strictly increasing in k_{t+1}, for values of k_t and i_t the value of k_{t+1} is determined. So let $P(\{k_{t+1}\}|k_t,i_t)=1$ if $i_t=g(k_t,k_{t+1})$ and zero otherwise.

The assumption of constant returns to scale places no restriction on the number of firms operating in the industry. Hence, without loss of generality we can consider the case where there is a continuum of firms operating in the industry at time zero with an initial distribution μ_0 of capital stocks[8]. Let $y_t=(q_t,i_t)$ and $p(\theta_t,y_t)=(D(\theta_t,q_t),1)$. With this

[8]This distribution can have point masses at the capital stocks that would correspond

specification, the model considered by Lucas and Prescott can be accommodated to the structure presented in this paper.

Jovanovic (1982) develops a model of an industry where each firm's variable cost function depends on the realization of a shock η_t drawn each period from a normal distribution with known variance but unknown mean θ. This mean is firm specific and time invariant. Entrants' value of θ are normally distributed according to a known distribution with mean $\bar{\theta}$. Based on the realization of the cost parameter η_t, firms update their prior beliefs, which are summarized by a prior mean and precision. More precisely, let n be the age of a firm and θ_n its prior mean; so $\theta_n = \sum_{i=1}^{n} \eta_i$ and $\theta_{n+1} = \frac{n\theta_n + \eta_n}{n+1}$.

Let the state of the firm be $s_n = (\eta_n, \theta_n, n)$ and the transition function defined by using the above law of motion for θ_n and the normal prior distribution for η_{n+1}, imposing that with probability one the last component will be n+1. The initial distribution is defined in an analogous way but taking $\bar{\theta}$ as initial prior mean.

In addition to the variable costs, firms pay a constant fixed cost per period and there is a sunk entry cost for new entrants. To include the cost of entry, we proceed in the following way. Let s be the state of an incumbent firm and s_e be identical to s except for a component that indicates the firm is an entrant. Then $Y(\theta_t, s_e) = Y(\theta_t, s) - \mathbf{c}$, where $\mathbf{c} = (0, 0, \ldots, 0_{k-1}, c)$ and $p_k(y) = 1$ for all y.

There is no capital in the model so firms make no investment decisions. Prices each period are determined competitively, where the aggregate demand function for each period is known and follows a deterministic process. The timing of entry, production and exit

---

industry with finitely many firms.

decisions is similar to the one described in the introduction.[9] The equilibrium concept employed by Jovanovic is analogous to the one given in this section.

With the above definition of the state vector, mapping the model into the general framework presented is immediate.

4. Equilibrium Existence and a Characterization.

This section describes the main approach followed in the existence proofs and provides conditions under which an equilibrium exists and the equilibrium prices are unique. We first discuss some distinctive features of the theory developed that motivate some of the assumptions we make (section 4.1). The main results are then presented in 4.2 , followed by a discussion of an algorithm that can be used to compute the equilibria based on the methods used in the proofs (4.3). The more technical matters are left to the appendix.

4.1 Assumptions.

The strategy of the proof is to show that the equilibrium allocations maximize the expected discounted sum of net consumer surplus on the set of feasible allocations (the 'planning problem') and, conversely, any allocation that solves this planning problem is an industry equilibrium –for the price process generated by the allocation. This is also the

[9]There is a slight difference since in Jovanovic(82) firms make their output decision prior to observing their cost parameter. So essentially firms precomit themselves to an output level; this can be handled by making one of the components of the state of the firm s_t be the 'output decision for the current period' and restricting the technology correspondence in the obvious way.

approach followed in Lucas and Prescott, 1972. Our environment is more complex in several respects, which are now discussed.

The first difference arises from the heterogeneity of firms and lack thereof of a 'representative' firm, as occurs in their framework. As a consequence, there is no simple summary statistic for the state of the system, as given in Lucas and Prescott by the aggregate capital stock, so the whole distribution of firms' states and its evolution must be taken into consideration. This not only complicates proving the existence of a solution to the programming problem, but introduces an additional complication in establishing the duality between industry equilibria and solutions to that program: instead of one representative firm maximizing profits, we have a measure space of them. However, these firms make decisions independently faced with a common stochastic process for prices and there are no externalities. It is well known that under these conditions and with a finite number of firms, profit maximizing by each firm separately gives the same aggregate allocations as the ones obtained by maximizing their joint profits. To establish this in our framework is not a trivial exercise.

Firms have been 'forced' (by our construct) to be small (nonatomic) relative to the size of the industry. This is obviously inconsistent with the existence of increasing returns beyond a limited scale. Moreover, having admitted the possibility of heterogeneity, it is apparent that if some firms in the industry are 'better' than others, production may tend to concentrate on the better ones. And unless some assumptions are made to limit the efficient scale and/or the relative advantage of firms, there may be no solution (and equilibrium) to our problem.[10] The restriction of the optimal size of firms to a 'small efficient scale' is

[10]Suppose the initial distribution is given by a density of firms over an interval that reflects their unit costs of production and that the state of firms does not change through time. The optimal production plan would entail assigning all production to the firm with the lowest

precisely the purpose of assumptions (SES) and (SES') below. The former imposes an upper limit to capacity, while the latter introduces asymptotic diseconomies of scale with respect to an essential input, which has strictly positive price. These assumptions are used alternatively in the existence proof (Proposition 4).

In contrast, if firms production technologies exhibit decreasing returns to scale and there are no costs of entry to the industry, the opposite problem arises: the mass of firms in the industry could become unbounded. This obviously translates into another non existence problem. There are two natural ways to avoid this: to rule out decreasing returns 'at zero', e.g. with a fixed cost, or to include setup costs for entrants. Our assumption (T3), which is consistent with both interpretations, heuristically states that "large entry requires large inputs".

One of the requirements in the definition of equilibrium is that (3.1) be well defined. We have already limited the efficient scale of firms, so provided that prices are adequately bounded, (3.1) will be well defined. Assumption (B3), which essentially says that $\beta^t p_t$ is bounded by an integrable function, provides this bound.

We now describe the 'joint profit maximization' problem mentioned above. Consider a price process $\{p_t\}$ and let $\{y_t\}$ be a feasible input output process such that $E_0 \Sigma \beta^t p_t y_t$ is

---

unit cost. But this firm has measure zero! It is easy to see that there will be no solution to the optimal program. Technically, this appears as a discontinuity of the aggregate input output vector with respect to the input output allocation $\mathcal{y}$.

well defined.[11] If $E_0\Sigma\beta^t p_t y_t$ is finite, we will say that the input ouput process has <u>bounded value</u>. The value maximization problem faced by the 'aggregate' firm can be stated as:

$$(4.1) \qquad \max_{\{y_t\}} E_0 \sum_{t=0}^{\infty} \beta^t p_t y_t$$

on the set of feasible allocations with bounded value.

We define a feasible allocation $\{y_t^*, M_t^*, e_t^*, \mu_t^*\}_{t=0}^{\infty}$ together with a price process $\{p_t^*\}$ as a <u>single firm equilibrium</u> if it satisfies (v) of the definition of industry equilibria and is a solution to the above problem. For this problem to be well defined, some assumptions that provide a bound on revenues will be made later.

To define the optimal program, for $y \in Y$ let

$$(4.2) \qquad S(\theta,y) = \sum_{\{j|y_j \geq 0\}} \int_0^{y_j} p_j(\theta,x)dx - \sum_{\{j|y_j<0\}} \int_{y_j}^{0} p_j(\theta,x)dx.$$

This objective has implicit the fact that the allocation pricing function has the form $p(y) = [p_1(y_1),p_2(y_2).,,,p_k(y_k)]$, so that there are no cross–price effects, which is assumption (D1). This has been done to guarantee that the gradient vector of $S(\theta,y)$ is $p(\theta,y)$ and that, provided each price is (weakly) decreasing in the corresponding aggregate, $S(\cdot)$ is concave in y. These are the key features needed to establish the duality between equilibria and solutions to the optimal program, so any alternative set of assumptions that gives a function $S(\theta,\cdot)$ with those features can be equally used. The planning problem is defined by:

---

[11]In all our definitions we consider the integral with respect to the product of the measure over sample paths and a counting measure on the positive integers. Thus $E_0\Sigma\beta^t p_t y_t$ is to be taken as the integral of the function $\beta^t p_t y_t$ with respect to that measure. Note however that as far as this integral exists, by Fubini's theorem the iterated integrals do too.

(4.3) $\max E_0 \sum_{t=0}^{\infty} \beta^t S(\theta_t, y_t)$

over the set of feasible allocations, where $y_t = \int y\, \gamma_t(ds,dy)$ and $S(\cdot)$ is defined by (4.2).

Heuristically, this is the problem faced by a planner that maximizes the expected discounted surplus generated by the industry subject to the technological constraints, by opening, operating and closing 'plants' every period. To assure this problem is well defined, we will assume that $S(\cdot)$ is bounded above (B1). Assumptions (SES) or (SES') limit the optimal size of firms. The 'size of the industry' needs to be limited too, for even if $S(\cdot)$ is bounded it can be strictly increasing in the input output vector y. Assumption (B2) provides this limit; it loosely states that 'large' input output vectors are undesirable.

We now list the main assumptions used in our proofs.

<u>Assumptions</u>.

(S1) S is a compact metric space.[12]
(S2) Θ is a borel subset of a metric space.
(S3) $P(ds',\cdot)$ is a weak* continuous function from $\Theta \times S \times Y$ to $\mathcal{P}(S)$, the probability measures on S.

---

[12]The compactness of S is not necessarily a significant restriction: for the endogenous components (e.g. capital) bounds can often be established, as indicated by some of the examples in section 6. For non compact exogenous processes it is sometimes possible to resort to a compactification argument, as done in example b) of Section 6.

(D1) $p(\theta^t,y)=\{p_j(\theta^t,y_j)\}_{j=1}^k$ (independent pricing functions).

(D2) p_j is measurable in $\Theta^t \otimes L$, where L corresponds to the lebesgue sets and $p(\theta^t,\cdot)$ is continuous for almost all θ^t.

(D3) $p(\theta^t,\cdot)$ is (weakly) decreasing.

This is well defined by assumption (D1), (D2) and (B3) below. Let $\mathcal{Y}$ denote the cone generated by Y (i.e. the set of all points in $\mathbb{R}^n$ of the form λy where $y \in Y$ and $\lambda \geq 0$).

(B1) $S(\cdot)$ is bounded above.

(B2) For any sequence $y^n \in \mathcal{Y}$ such that $\|y^n\| \to \infty$ and $\theta^t \in \Theta^t$, $\limsup S(\theta^t,y^n)=-\infty$.

(B3) $p(\cdot)$ is bounded by a β–integrable function.[13]

(B3') $p(\cdot)$ is uniformly bounded above.

(T1) $Y(\cdot)$ has closed graph in $\theta \times S \times Y$.

(T2) For any initial measure μ_0 there exists a feasible allocation.

(T3) $M_t \to \infty$ implies $\|y_t\| \to \infty$ or $M_t \leq \overline{M}_t < \infty$.

(SES) Y is bounded above. (recall $Y(\theta^t,s) \subset Y$).

(SES') (a) For every $\epsilon > 0$, $y \in Y$ and $\|y\| \geq c$ imply $\frac{\|y\|}{|y_j|} < \epsilon$, where c only depends on ϵ and $y_j < 0$, where j is fixed.

(b) The price of input j is strictly positive, i.e. $p_j(\theta^t) \geq \underline{p}_j > 0$.

[13]The function X_t is β–integrable if $\beta^t X_t$, viewed as a function of (t,θ) is integrable with respect to the product of measure λ and a counting measure on the positive integers.

4.2. Existence and Uniqueness of Equilibrium.

The proof of existence is done in the following steps: First, it is established that the allocations that maximize the expected discounted sum of net consumer surplus on the set of feasible allocations (the 'planning problem') provide a single firm equilibrium –for the price process generated by these allocations– and that any single firm equilibrium is also a solution to the planning problem. These results are summarized in Propositions 1 and 2, which are proved using very simple variational arguments. Secondly, it is shown that for any price process, feasible allocations that solve (4.1) satisfy conditions (i)–(iv) of the definition of Industry Equilibrium, and conversely, any allocations that satisfy (i)–(iv) are a solution to (4.1). These results are summarized in Proposition 3. Finally, a solution to the planning problem (4.2), and thus the existence of equilibria, is established in Proposition 4. These results are summarized in the main existence result, Theorem 1. Since not all assumptions are necessary for each step, we will follow the convention of including at the beginning of each proposition and within brackets, the assumptions used So, for example, (S) without affixes represents assumptions (S1)–(S4) and similar conventions are used for the rest of the assumptions.

Proposition 1. (S)(D)(B1) Suppose $\{p_t^*, \mathcal{Y}_t^*, M_t^*, e_t^*, \mu_t^*\}_{t=0}^{\infty}$ is a single firm equilibrium. Then $\{y_t^*, M_t^*, e_t^*, \mu_t^*\}_{t=0}^{\infty}$ is a solution to the optimal program.

Proof. Let $\{y_t\}$ be the aggregate production plan corresponding to $\{\mathcal{Y}_t\}$. By Lemma 1, $E_0\Sigma\beta^t S_t(\theta^t, y_t)$ exists and exceeds $-\infty$. Suppose by way of contradiction that there exists another feasible input output process $\{y_t'\}$ such that $E_0\Sigma\beta^t S(\theta^t, y_t') - E_0\Sigma\beta^t S(\theta^t, y_t) > 0$. Let $\{y_t^\delta\}$ be the process defined by $y_t^\delta = \delta y_t' + (1-\delta) y_t$. For notational convenience let $p_t = p(\theta^t, y_t)$ and denote by S_t, S_t' and S_t^δ the function S evaluated at θ^t and, respectively, at y_t, y_t' and y_t^δ.

Since the set of feasible allocations is convex, $\{y_t^\delta\}$ is feasible for all $\delta\in[0,1]$. By Lemma 2, $(S_t^\delta-S_t)/\delta$ is decreasing in δ and by Lemma 1 it increases to $p(\theta^t,y_t)(y_t'-y_t)$ as $\delta\downarrow 0$. Since for $\delta\in(0,1]$, $(S_t^\delta-S_t)/\delta$ is bounded below by $S_t'-S_t$, which is an integrable function, so using the monotone convergence theorem

$$
\begin{aligned}
0 &< E_0\Sigma\beta^t S_t' - E_0\Sigma\beta^t S_t \\
&\leq \lim_{\delta\downarrow 0}\, [E_0\Sigma\beta^t S_t^\delta - E_0\Sigma\beta^t S_t]/\delta \\
&= \lim_{\delta\downarrow 0}\, E_0\Sigma[\beta^t(S_t^\delta - S_t)/\delta] \\
&= E_0\Sigma\beta^t[p_t(y_t'-y_t)] \\
&= E_0\Sigma\beta^t p_t y_t' - E_0\Sigma\beta^t p_t y_t
\end{aligned}
$$

a contradiction to the fact that y_t is a profit maximizing process for $\{p_t\}$ □.

Remark. A sufficient assumption to justify the exchange of limits and interchange in the order of integration is that $S(\cdot)$ be bounded above by a β–integrable function.

Proposition 2. (S)(D)(B1,B3) Suppose $\{\mathcal{y}_t^*,M_t^*,c_t^*,\mu_t^*\}_{t=0}^{\infty}$ is a solution to the optimal program. Then this allocation together with prices $p_t^*=p(\theta^t,y_t)$, where $y_t=\int y\mathcal{y}(ds,dy)$, is a single firm equilibrium.

Proof. Consider first an alternative bounded[14] input output process $\{y_t'\}$; note that by (B3) $E_0\Sigma p_t y_t'$ is finite and $E_0\Sigma\beta^t S_t'>-\infty$ and define y_t^δ as in the proof of the previous Proposition.

[14]We will say that the process $\{y_t\}$ is bounded if $\sup_{\theta^t,t} \|y_t(\theta_t(\omega))\|<\infty$.

By an argument similar to the one used in the previous proof,

$$0 \geq [E_0\Sigma\beta^t S_t^\delta - E_0\Sigma\beta^t S_t]/\delta \rightarrow E_0\Sigma\beta^t p_t(y_t' - y_t) \text{ as } \delta\downarrow 0.$$

Since $E_0\Sigma\beta^t p_t y_t' > -\infty$, the above implies $E_0\Sigma\beta^t p_t y_t' \leq E_0\Sigma\beta^t p_t y_t$ and by Lemma 1, $E_0\Sigma\beta^t p_t y_t \leq E_0\Sigma\beta^t S_t < \infty$ so $\{y_t\}$ has bounded value.

Consider now an arbitrary feasible aggregate input output process $\{y_t'\}$ with bounded value. For each path $\{\theta_t\}$ let $\tau_n = \sup\{t \mid \|y_t\| \leq n\}$. Let $y_t^n = y_t'$ for $t \leq \tau_n$ and zero otherwise. Note that this plan is feasible since all firms can exit after observing θ^t. Since $\{y_t'\}$ has finite value, $E_0\Sigma\beta^t p_t y_t^n$ is finite. Let $A_n = \{\{\theta_t\}, t \mid \tau_n(\{\theta_t\}) < t\}$. Since $A_n \downarrow \phi$,

$$E_0\Sigma\beta^t p_t y_t^n = E_0\Sigma\beta^t p_t y_t - E_0\Sigma\beta^t \chi_{A_n} p_t y_t \rightarrow E_0\Sigma\beta^t p_t y_t$$

which proves the claim □.

The following proposition states the equivalence between single firm equilibria and industry equilibria.

Proposition 3. (S)(D)(B1–B3')(T)(SES or SES') The set of SFE and Industry equilibrium coincide.

Proof. See appendix.

Remark. Assumption (B3') allows us to use some results in negative dynamic programming. It is an open question whether the same can be done under (B3), which we conjecture is true.

Proposition 4. If assumptions (S)(D)(B1–B3)(T) and either (SES) or (SES') are satisfied, there exists a solution to the optimal program.

Proof. See appendix.

We now state one of the main results, a direct consequence of Propositions 1 thru 4

Theorem 1. If assumptions (S)(D)(B1–B3')(T) and either (SES) or (SES') are satisfied, there exists an Industry equilibrium.

Having established the existence of equilibria we now turn to the question of uniqueness. The following Theorem states (loosely) that equilibrium prices are unique.

Theorem 2. Let p_t and p_t' be two equilibrium price processes. Then $p_t = p_t'$ (almost everywhere).

Proof. If for some θ^t, y and y' are two input output vectors and $p_j(\theta^t, y) \neq p_j(\theta^t, y')$, then for any $\lambda \in (0,1)$, $S(\theta^t, y^\lambda) > \lambda S(\theta^t, y) + (1-\lambda) S(\theta^t, y')$, where $y^\lambda = \lambda y + (1-\lambda) y'$. In words, $S(\theta^t, \cdot)$ is strictly concave over aggregate input output vectors with different prices. This fact, together with the convexity of the set of feasible allocations implies that any two distinct equilibrium allocations must (almost everywhere) have the same prices □.

Corollary. If for some j, $p_j(\theta^t, .)$ is strictly decreasing for all θ^t, then for all equilibrium allocations the process for y_{jt} coincides □.

It is worth emphasizing that in many cases the above two results can be used to establish uniqueness of the equilibrium. The following case, studied extensively in Hopenhayn (1989), indicates this connection. Suppose that S is equal to an interval $[s_1,s_2]$ in the real line, that for any price vector p the profits of a firm $\pi(s',p)>\pi(s,p)$ whenever $s'>s$ and that for any price process the input output choices of firms are uniquely determined. Note that this implies that all firms with state below certain threshold $s^*(\theta^t)$ exit, while all firms above that threshold remain in the industry. For simplicity assume that the distribution function for firm's states is continuous.[15] Let one of the goods be an output good, and assume that in any equilibrium its price and the quantity produced by each firm are strictly positive. Suppose the price of this good is strictly decreasing in the aggregate quantity demanded. Under these conditions, by the above corollary the total output process for this good is uniquely determined. We now show that given a continuous initial distribution, the equilibrium is unique. To see this, note that since total output is strictly increasing in the number of entrants, entry for the first period is uniquely determined. With the above exit rule and a continuous distribution of firms' states, given θ^1 the exit of firms is unambiguously determined. But this implies that the initial distribution for the following period, namely μ_1, is also uniquely determined. Repeating the argument, given (θ_1,θ_2) M_2 is determined. This argument, applied recursively, implies there is a unique equilibrium.

4.3. Dynamic Program: A Decomposition.

We have shown that the equilibrium allocations correspond to solutions to a dynamic programming problem. For the special case where the transition function is independent of y, this problem can be decomposed into a 'static' and a 'dynamic' part. The former

[15]This will be the case if for all (θ_t,s,y) the conditional distribution $F(ds'|\theta_t,s,y)$ and the initial distribution $\nu(ds'|\theta^t)$ are continuous.

corresponds to the optimal input output choice, which only depends on the distribution of firms active in a given period and the aggregate shock. The 'dynamic' part corresponds to the choices of mass of entrants and distribution of exits. Let $R(\theta_t,\tilde{\mu})$ denote the maximum surplus attainable in the current period when the distribution (after entry) is $\tilde{\mu}$. The functional equation simplifies to:

$$V(\theta^t,\mu) = \max_{\{M,e\}} R(\theta_t\,,\,\mu+M\nu_t)+\beta E_t V(\theta^{t+1},\mu')$$

$$\text{subject to: (i) } M\geq 0$$
$$\text{(ii) } 0\leq e\leq \mu$$
$$\text{(iii) } \mu'(ds')=\int P(ds'|\theta^t,s)[\mu(ds)+M\nu(ds|\theta^t)-e(ds)]$$

This is a dynamic programming problem with a nonlinear objective but linear constraints, which from the computational point of view, makes the evaluation of $V(\theta^t,\mu)$ for each iteration on the above bellman equation a relatively simple task. For illustrative purposes, we will comment on a specific application.

The author has used this algorithm to study the adjustment of an industry to a demand change, the evolution of an industry faced with a (deterministic) demand cycle and the effect of exogenous technological change on industry dynamics. The model used has a discrete aggregate process $\{\theta^t\}$ and a finite number of idiosyncratic states for the firms $(s_1,\ldots,s_n)$. To simplify the exposition, however, we only include the firm specific process. The initial distribution is given by an $n\times 1$ density vector **v** and the transition function –only dependent on the current state– is given by an $n\times n$ matrix **P**, where element p_{ij} denotes the probability of going from state s_i to states s_j. Since the choice of output does not affect the transition of firms, the planning problem can be decomposed in the manner described above. The production function is given by $f(s,n)=s[\gamma_1 n-\gamma_2 n^2]$, where n is the only input, the inverse demand function linear (with constant A and slope b) and the price of the input fixed at 1. The surplus function is therefore $S(Q,N) = AQ - \frac{b}{2}Q^2 - N$. All firms in the industry incur in fixed cost c_f and entrants have an entry cost c_e. If the distribution of firms available for production in a period is given by the vector $\bar{\mathbf{u}}$, where element $\bar{u}_j$ indicates the mass of firms with state s_j, the problem faced by the planner at the production stage is:

$$R(\bar{\mathbf{u}}) = \max_{n_j} AQ - \frac{b}{2}Q^2 - N - c_f\sum_j \bar{u}_j$$
$$\text{where } Q=\sum_j \bar{u}_j s_j[\gamma_1 n_j - \gamma_2 n_j^2] \text{ and } N=\sum_j \bar{u}_j n_j.$$

The recursive problem is given by:

$$V(\mathbf{u}) = \max_{M,\mathbf{e}} R(\mathbf{u}+M\mathbf{v}) - c_e M + \beta V(P[\mathbf{u}+M\,\mathbf{v}-\mathbf{e}])$$
$$\text{subject to: } M \geq 0 \text{ and } 0 \leq e_j \leq u_j + Mv_j,$$

where $\mathbf{e}=(e_1,\dots,e_j)$ is the vector of exits.

The algorithm can be implemented by solving this recursive problem or truncations of the corresponding sequential problem.

5. <u>The Stationary Model</u>

An interesting question is under what conditions an equilibrium process has a stationary distribution. This is particularly useful for statistical analysis, e.g. to obtain consistent estimates of parameters of the equilibrium process from the sample averages.[16] This is related to a more general question which is under what conditions the equilibrium process varies continuously with initial conditions or with parameters of the model. The following section gives assumptions under which the optimal policy rule is an upper hemicontinuous function of the state. This is used to establish that there exists a stationary distribution for the equilibrium process, provided the equilibrium satisfies a boundedness condition.[17] We now present the additional assumptions.

(M1) The process θ_t is first order markov with continuous transition function and takes values on a compact metric space Θ.

(D2') $p(\cdot)$ is jointly continuous.

(T3') $Y(\cdot)$ is a continuous correspondence.

Finally, while convexity of the individual firms production sets was not required in the general case, it will be used here to get the desired continuity results. Hence we assume

(M2) For each $\theta \epsilon \Theta$ the graph of $Y(\theta, \cdot)$ is convex.

An industry satisfies the <u>stationary model</u> if assumptions (S)(D)(B)(T)(SES) or (SES') and additionaly (D2')(T3')(M1)(M2), hold.

---

[16]See Duffie et al. for more on this.

[17]The methods used to prove the continuity of the optimal policy rule can also be applied to establish that the equilibrium process changes continuosly with the value of parameters of the model.

As will become clear in the proof of Proposition 5, the optimal program in the stationary model is generated by the unique fixed point of the following functional equation on the space $C(\Theta,S)$ of continuous and bounded functions on $\Theta\times S$.

$$(5.1)\qquad V^{k+1}(\theta,\mu) = \max_{\{M,\mathcal{y},e\}} S(\theta,y) + \beta\int V^k(\theta',\mu'|\theta)$$

subject to $(\mu', M, y, \mathcal{y}, e)\in\Gamma(\theta,\mu)$

where $\Gamma(\theta,\mu)$ is defined by:

1. $y=\int y\,\mathcal{y}(ds,dy)$.
2. $\mathcal{y}$ has first marginal $\mu+M\nu_\theta$ and support on the graph of $Y(\theta,\cdot)$.
3. $0\leq e\leq \mathcal{y}$.
4. $\mu'(ds') = \int P(ds'|\theta,s,y)[\mathcal{y} - e](dy,ds)$.

Let $g(\theta,\mu)$ be the stationary optimal rule that solves the optimal program, described in the proof of Proposition 4.

Proposition 5. The optimal decision rule of a stationary model is upper hemicontinuous.

Proof. The optimal program defines a discounted dynamic programming model and any fixed point of (5.1) gives the value under the optimal policy. We will first establish that the mapping defined by (5.1) is a contraction mapping on the space of continuous and bounded functions on $(\Theta\times S)$.

It is easy to verify that function $S(\cdot)$ is continuous and by (B1) bounded above. Assumption (B2) implies that, without loss of generality, the aggregate input output vector can be chosen so that $\|y\|\leq b$ for some $b<\infty$. The set of measures $\mathscr{y}$ that satisfy this restriction are a tight class and by (T3) M is also bounded above. This implies that Γ is compact valued. Assumption (SES) implies that $\mathscr{y}$ has support on a compact set. Alternatively, under assumption (SES') and (B3) revenues are uniformly bounded above across all equilibria. Since firms have the possibility of exit, this also implies that in equilibrium they will be uniformly bounded below. This implies that the support of $\mathscr{y}$ can also be restricted to a compact set. With this restriction, it is easy to verify that the input output vector y is continuous in $\mathscr{y}$. Since $P(ds';\theta,s,y)$ is continuous, and $\Theta\times S$ compact, the equation defined in 4. is continuous in $(\mathscr{y},e,\theta)$. We now show that condition 2. has a closed graph. Let $\theta_n\rightarrow\theta$, $\mu_n\rightarrow\mu$ and $\mathscr{y}_n\rightarrow\mathscr{y}$, where $\mathscr{y}_n$ have first marginal μ_n and support on gr $Y(\theta_n,\cdot)$. Since $\mathscr{y}$ is a regular measure, given $\epsilon>0$ there exists a set U' such that $\mathscr{y}(U')-\mathscr{y}(\Lambda)<\epsilon$, where Λ is the graph of $Y(\theta,\cdot)$. Choose U open and C closed sets such that $\Gamma\subset U\subset C\subset U'$. Since $Y(\cdot)$ is upper hemicontinuous, for θ_n on a neighborhood of θ $gr(Y(\theta_n,\cdot)\subset U$ so

$$\mathscr{y}(S\times Y) = \limsup \mathscr{y}_n(C) \leq \mathscr{y}(C) \leq \mathscr{y}(\Gamma)+\epsilon$$

so $\mathscr{y}$ has support on the graph of $Y(\theta,\cdot)$. That the first marginal is μ follows from weak convergence. Finally condition 3. can be easily checked to be a closed relation, so Γ is upper hemicontinuous.

We will now prove that it is lower hemicontinuous too. First we will show that given sequences $\mu_n\rightarrow\mu$ and $\theta_n\rightarrow\theta$ and $\mathscr{y}$ with first marginal μ and support on gr $Y(\theta,\cdot)$, there is a sequence $\mathscr{y}_n\rightarrow\mathscr{y}$ with first marginals μ_n and support on gr $Y(\theta_n,\cdot)$. By Lemma 3 there is a sequence $\mathscr{y}'_n\rightarrow\mathscr{y}$ with first marginals μ_n. We will 'distort' those measures appropriately so that they satisfy the correct support restriction. Since $\Theta\times S$ is compact, Y a subset of a banach space and $Y(\theta,s)$ convex, by lemma [ ] there exists a continuous function $\omega:\Theta\times S\times Y\to Y$ such that $\omega(\theta,s,y)\in Y(\theta,s)$ and $\omega(\theta,s,y)=y$ whenever $y\in Y(\theta,s)$. Let $h(\theta,s,y) = (\theta,s,\omega(\theta,s,y))$. Define $\mathcal{y}_n$ by $\mathcal{y}_n(B) = \int \chi_B(h(\theta_n,s,y))\mathcal{y}'_n(ds,dy)$. With this definition, $\mathcal{y}_n(A\times Y)=\mu_n(A)$ and for any continuous and bounded function $f:S\times Y\to\mathbb{R}$

$$\int f(s,y)\mathcal{y}_n(ds,dy) = \int f(h(\theta_n,s,y))\mathcal{y}'_n(ds,dy) \to \int f(h(\theta_n,s,y))\mathcal{y}(ds,dy) = \int f(s,y))\mathcal{y}(ds,dy).$$

Thus the sequence $\mathcal{y}_n$ satisfies all the requirements.

Now let $\theta_n\to\theta$ and $\mu_n\to\theta$ and suppose $(\mu', M, y, \mathcal{y}, e)\in\Gamma(\theta,\mu)$. Let $M_n=M$ and choose a sequence $\mathcal{y}_n$ with marginal $\mu_n(ds)+M\nu(ds|\theta_n)$ and support on gr $Y(\theta_n,\cdot)$ that converges to $\mathcal{y}$. By lemma [ ] there exists a sequence $e_n \to e$ such that $0 \leq e_n \leq \mathcal{y}_n$. Given these choices, 1. and 4. define sequences y_n and μ'_n that converge, respectively to y and μ. So Γ is lower hemicontinuous.

Using the theorem of the maximum, equation (5.1) defines a mapping from continuous and bounded functions into itself. This mapping satisfies the monotonicity and discounting assumptions in Blackwell, and hence it is a contraction. Its unique fixed point gives the value of the optimal program and, again using the theorem of the maximum, an upper hemicontinuous optimal policy function □.

With this proposition, it is now easy to show that there exists a stationary distribution for the markov process generated by the decision rule g, provided the equilibrium allocations are bounded, as defined in the following assumption.

(M3) There is a closed ergodic subset Θ' of Θ and $b>0$ such that $\mu_0(S)\leq b$ and $\theta_1\in\Theta'$ imply $\mu_t(S)\leq b$ for any equilibrium.

Theorem 3. A stationary model that satisfies (M3) has an Ergodic Markov Equilibria.

Proof. Proposition 5 implies that g is an expectations correspondence. Assumption (M3) implies that $\Theta'\times\{\mu\in\mathcal{M}(S)|\mu(S)\leq b\}$ is self justified. By Proposition 1.2 in Duffie, Genakoplos, Mas–Colell and McLennan, there exists an Ergodic Markov Equilibrium □.

6. Applications.

In this section we present applications of the general theory to several models.

a. Investment Under Uncertainty. The model has been described in section 3. Lucas and Prescott assume that the price of the homogeneous output is $D(\theta^t,q_t)<\bar{\bar{p}}<\infty$. Given the assumptions on the adjustment cost function, it is not hard to show that any equilibrium must have $k_t\leq\bar{k}$ for some $\bar{k}<\infty$. Thus without loss of generality we can assume the state space is compact. Since (S2) is satisfied and (S3) is an immediate consequence of the continuity of the function $g(\cdot)$, the first three assumptions are satisfied. Assumptions (D1)–(D3) are immediately verified. Consumer surplus is assumed uniformly bounded above, and since the price of the investment good is one, (B2) holds. We have already mentioned that the output price is bounded, so (B3) is satisfied. (T1) and (T2) are immediate. Entry is excluded from the model, so (T3) is immediately satisfied. Choosing $Y_0=\{\bar{k},0\}$, (T4) is satisfied and (SES) holds since Y is bounded above by the same vector. Hence the assumptions of Theorem 1 are satisfied.

We now show that there is a unique equilibrium. Inverse demand is strictly decreasing so aggregate output must coincide for all equilibria. Given initial strictly positive capital stocks for the firms, the problem (3.1) faced by the firm is strictly concave in the choice of capital accumulation paths. Thus the investment policy of firms is the same throughout all equilibria. The only source of indeterminacy that may arise is in periods where demand realizations are too low for all capital to be fully employed. In this case, while total output will be determined in equilibrium, its distribution across different firms will not, and obviously firms will be indifferent between producing or not.

Assuming the process for the aggregate demand shock is first order markov with a continuous transition function and takes values on a compact set, theorem 3 applies and there exists a stationary distribution for the equilibrium process.

b. <u>Jovanovic's Selection Model</u>. The model has been broadly described in section 3.2 , but to check that the assumptions are satisfied, more details need to be provided. The variable cost of production of the firm is $c(q_t)x_t$, where $x_t=\xi(\eta_t)$, $\xi(\cdot)$ is a positive, strictly increasing and continuous with $\lim_{\eta_t\to-\infty}\xi(\eta_t)=\alpha_1>0$ and $\lim_{\eta_t\to\infty}\xi(\eta_t)=\alpha_2\leq\infty$. Since this function is strictly increasing, there is a one–one correspondence between values of $x\epsilon[\alpha_1,\alpha_2]$ and $\eta\epsilon R$. Jovanovic shows that (x_t^*,n_t) is also a sufficient statistic, where x_t^* is the expected value of x_{t+1} with respect to its prior distribution. Define $s_t=(x_t,x_t^*,n_t)$ as the state of the firm at t.[18]

[18]To make the state space compact, the endpoints α_1 and α_2 must be included as well as $n=\infty$. The transition function needs to be defined at these points too. Note that for x_t^* to be very close to α_1 or α_2 the variance of the prior for x_{t+1} must be very small. So the natural thing to do is to set this variance at zero (or precision ∞) for x_t^* at α_1 or α_2. Since the precision of the prior approaches ∞ as $n\to\infty$, the same will be assumed for $n=\infty$. Also note that the posterior after observing $x_t=\alpha_1$ or α_2 must put probability one on these values. It can be checked that with these definitions the transition function will be continuous.

Let the input output vector be given by (q,z), where q is the output of the firm and z will be an input associated to the entry cost (c_e), fixed cost (F>0) and variable costs, with price normalized to one. For an incumbent firm with state $s_t=(x_t,x_t^*,n)$, $Y(s_t)=\{q,z: z\geq F+c(q)\, x_t\}$, and for an entrant with $s_t=(x_t,x_t,1)$, $Y(s_t)=\{q,z: z\geq c_e+F+c(q)x_t\}$.[19] There is no aggregate uncertainty in Jovanovic's model, although he allows the inverse demand function $D_t(\cdot)$ to be time dependent. To accommodate that, we define the aggregate process $\theta_t=t$. The allocation pricing function will be given by $p_1(t,q,z)=D_t(q)$ while $p_2=1$ for all allocations.

Jovanovic assumes there is a uniform upper bound to the consumer surplus so (B1) is satisfied. If $q_t\to\infty$ then either the number of firms approaches ∞ or the output of a set of firms with positive measure does so. Given that each firm has a fixed cost F>0, in the former case $z_t\to\infty$. Since variable costs are bounded below by $c(q)\alpha_1$ which by the assumptions is strictly positive, increasing and convex, the latter case also implies $z_t\to\infty$, so (B2) is satisfied. For sufficiently high price ($\bar{p}$) entry would be profitable even if firms produced for just one period. Though Jovanovic does not assume that $D_t(\cdot)$ is bounded above, there is obviously no loss of generality in redefining it to be $\min\{\bar{p},D_t(q)\}$ so that (B3') is satisfied. Assumptions (T1)–(T3) are easily verified. Jovanovic assumes $c'(q)\to\infty$ as $q\to\infty$, which implies that $q/(c(q)\to 0$. Since the cost function $c(q)x\geq c(q)\alpha_1$ and $\alpha_1>0$, condition (SES') is satisfied. So this model satisfies the assumptions of Theorem 1.

The inverse demand function is strictly decreasing, so by the corollary to Theorem 2 aggregate output must coincide for all equilibria. Furthermore, the exit rule (firms exit

[19] To be more rigourous, a dummy element indicating whether the firm is an entrant or not should be added to the state vector.

whenever their prior mean is above a threshold which may vary through time) is the same for all equilibria. Applying the same arguments as given in the remarks following the corollary to Theorem 2, it is easy to establish that the equilibrium is unique.

c. Lambson's model (identical firms) Lambson (1988) develops an industry equilibrium model where all firms are identical. There is a single output with cost of production given by an increasing and convex function $c(\theta_t,q_t)$, where θ_t is the realization of a stochastic process common to all firms. The aggregate process has countably many possible realizations. There is a fixed cost of operation $\varphi(\theta_t)$, an entry cost $\xi(\theta_t)$ and firms that exit receive a 'scrap value' $\chi(\theta_t)$, where $\xi(\theta_t)\geq\chi(\theta_t)$. Given the aggregate output of the industry, prices are determined by a continuous and decreasing inverse demand function $p(\theta_t,q_t)$. Output is uniformly bounded, i.e. for all firms $q_t\leq\bar{q}$ and $\lim_{q\to\infty} p(\theta,q)\to 0$.

There are two ways in which the 'scrap' value may be incorporated in our theory. The first is to include an element in the input output vector, say 'scrap', with price given by $\chi(\theta_t)$. Any incumbent firm can produce either one unit or zero of it. When a firm produces one unit, it moves in the following period to an absorbing state that makes the firm improductive. Obviously to avoid the fixed cost, it would exit at the end of the period. Alternatively, it can be shown that by properly modifying the fixed cost and cost of entry an industry $\mathcal{J}'$ with zero scrap value is obtained, and the equilibrium prices and allocations for $\mathcal{J}'$ coincide with those of the original industry.[20]

[20]Let v_t be the value of an incumbent firm in period t. Then $v_t=\pi_t + \max\{\chi_t, \beta E_t v_{t+1}\}$, so $v_t-\chi_t = \pi_t -\chi_t-(1-\beta)\chi_t + \max\{0,\beta E_t(v_{t+1}-\chi_{t+1}) + \beta(E_t\chi_{t+1}-\chi_t)\}$.
Let $\hat{v}_t\equiv v_t-\chi_t + E_{t-1}\chi_t-\chi_{t-1}$. Then $\hat{v}_t=\pi_t-\chi_t-(1-\beta)\chi_t+E_{t-1}\chi_t-\chi_{t-1}+\max\{0,\beta E_t\hat{v}_{t+1}\}$. By construction, $\chi_t>\beta E_t v_{t+1}$ if and only if $0>\beta E_t\hat{v}_{t+1}$, so as far as the price of the output good does not change, the exit rule in both industries is the same. To obtain the same entry rule

Though Lambson's boundedness assumptions are somewhat weaker than the ones used here, it is worth exploring the model with the boundedness restrictions given in our assumptions since this will allows us to see how the planning problem discussed in Section 4 can be used to characterize the equilibrium. This planning problem specializes to the following dynamic program:

$$V(\theta^t,z_t) = \max_{\{m_t,e_t\}} D(\theta_t,z_{t+1}) - m_t\xi_t - \varphi_t(\theta_t)z_{t+1} + e_t\chi_t + EV(\theta^{t+1},z_{t+1}|\theta^t)$$

with the restriction that $m_t \geq 0$ and $0 \leq e_t \leq z_t$, where $D(\cdot)$ is consumer surplus, z_t is the mass of firms as of the beginning of period t, m_t the mass entrants and e_t of exits and $z_{t+1}=z_t+m_t-e_t$. Since $\xi_t \geq \chi_t$, the solution to the above problem implies that whenever $m_t>0$, $e_t=0$ and the converse too. It is also easy to check that since $S(\theta_t,\cdot)$ is concave, $V(\theta^t,\cdot)$ is too. Hence we can conclude, as in Lambson, that the solution is characterized by stochastic boundaries $[N(\theta^t),X(\theta^t)]$ and that $z_{t+1}=z_t$ if z_t is in this region, $z_{t+1}=N(\theta^t)$ if z_t is below that point and $z_{t+1}=X(\theta^t)$ if it is above that point. Finally note that in the case where θ_t can take only finite number of values, the above recursive problem provides a simple algorithm for calculating these boundaries.

d. A model with heterogeneity and investment. Ericson and Pakes (1989) have recently developed an interesting model of industry dynamics which allows for investment and can generate entry, exit and heterogeneity in the size and growth rates of firms. For illustrative purposes, we focus on a simple case considered by them. Firms are distinguished by an integer valued state $\omega \epsilon N$. From one period to the next, this state can either remain

---

the entry cost needs to be adjusted accordingly, by letting $\hat{\xi}_t=\xi_t+\chi_t-E_{t-1}\chi_t+\chi_{t-1}$.

unchanged, move up one step or down one step. The outcome depends on the amount of investment made by the firm, denoted by $x \in \mathbb{R}_+$, and an aggregate shock. More precisely, the net change is given by:

$$p(1|x) = \begin{cases} p(1|x) = \lambda(x)(1-\delta) \\ p(0|x) = \lambda(x)\delta + [1-\lambda(x)](1-\delta) \\ p(-1|x) = [1-\lambda(x)]\delta \end{cases}$$

There are two possible realizations for the aggregate shock: a 'bad' (1) and a 'good' (0) shock, with probabilities δ and $(1-\delta)$, respectively. The bad shock moves the state of the firm down one step while the good one has no effect on it. On the other hand, the intrinsic luck of firms is represented by the function λ, which gives the probability that the firm moves up one step as a result of its investment x, with constant unit cost c. The net result is the sum of this aggregate and intrinsic shocks, as described above.

Profits of firms are given by a profit function, $\pi(\omega_t, y_t)$ which depends on the state of the firm and the number of 'active' firms in the industry in the period. A firm is active whenever its state $\omega \geq \hat{\omega}$. There is also a fixed (opportunity) cost c_f which all firms (active or not) pay each period and a fixed discount rate β. Firms maximize the expected discounted sum of net profits. In the special case considered the profit function has the form

$$\pi(\omega, y) = \begin{cases} A(y) \text{ if } \omega \geq \hat{\omega} \\ 0 \text{ otherwise} \end{cases}$$

where $A(\cdot)$ is a strictly decreasing function and $A(0)=A(1)=A<\infty$.

There is a cost of entry to the industry c_e. All entrants have initial state ω_0, but produce after the aggregate shock is realized. Hence their relevant state is ω_0 when the good aggregate state occurs and ω_0-1 otherwise.

Ericson and Pakes look at the case where there is a finite number of firms in the industry (rather than a continuum) and consider markov equilibria of the associated (anonymous) dynamic game. This, of course, is a valuable approach since firms have (and recognize having) an impact on the aggregate state of the industry. However, for an industry with a large number of firms, e.g. one where costs of entry and fixed costs are small, the relative impact of individual firms' actions on the aggregate state will be small, and aggregate uncertainty will come primarily from the aggregate shock. This makes a model with a continuum of firms a useful benchmark to consider. We now show how this can be done in our framework.

Let the aggregate state $\theta_t \in \{1,0\}$. We will denote the state of the firm by $s_t=\omega_t+\theta_t$. The production technology will be given by:

$$Y(\theta,s) = \begin{cases} (q,-\phi,-x),\ x\geq 0,\ x\geq 0 & \text{if } s-\theta\geq\omega \\ (0,-\phi,-x),\ x\geq 0 & \text{otherwise} \end{cases}$$

and prices given by $p_1(Q)=A(Q)$, $p_2=1$ and $p_3=c$.

Notice that this implies the indirect profit function given above.

Let $P(\{s-\theta+1\}|\theta,s,x)=\lambda(x)$ and $P(\{s-\theta\}|\theta,s,x)=1-\lambda(x)$.

To include the cost of entry, we can follow the procedure described in Section 2.

Using some additional assumptions Ericson and Pakes establish that all firms with state above a threshold will not invest, and that all firms with states below a certain threshold will leave the industry. The expected discounted returns of firms are uniformly bounded above by $\frac{A}{1-\beta}$, so the optimal investment policy has an upper bound. Hence S can be restricted to a compact set and so can Y. All remaining assumptions of our theorems are easily verified, so an equilibrium exists. Furthermore, since demand is strictly decreasing, the equilibrium price process is unique. Again, using the same arguments as in the remarks following the Corollary to Theorem 2, it is easy to show the equilibrium is unique. Finally, it is not hard to check that the assumptions of Theorem 3 are also satisfied, so the equilibrium process has a stationary distribution.

Mathematical Appendix

Lemma 1. i) S is measurable and differentiable in y. ii) S is concave in y and satisfies $\Sigma p_j y_j \leq S(\theta,y)$

Proof. i) It is enough to establish these properties for the functions S_j defined by $S_j(\theta,y) = \chi_{y_j \geq 0} \int_0^{y_j} p_j(\theta,x)dx - \chi_{y_j<0} \int_{y_j}^{0} p_j(\theta,x)dx$. Differentiability is immediate (even at y=0). Measurability is proved by applying (part of) Fubini's theorem (see Ash, Theorem 2.6.4) in the following way. Let $\omega_1=(\theta,y)$. For any lebesgue set $A \subset R$, Let $\mu(\omega_1,A)=\ell(A \cap [0,y])+\ell(A \cap [y,0])$, where ℓ denotes the lebesgue measure in the real line. Note that $\mu(\cdot,A)$ is a continuous function and hence measurable. It is also easy to see that μ is uniformly σ–finite. Then $S_j(\omega_1)=\int p_j(\theta(\omega_1),\omega_2)\mu(\omega_1,d\omega_2)$ is a measurable function.

(ii) The inequality and concavity are immediate from the fact that all p_j functions are nonincreasing (assumption D3) □.

Lemma 2. Consider two feasible input output processes $y' \in Y$ and $y \in Y$ and for $\delta \in [0,1]$ let $y^\delta=\delta y'+(1-\delta)y$. Then $[S(y^\delta,\theta)-S(y,\theta)]/\delta$ is decreasing in δ.

Proof. Let $g(\delta)=S(y^\delta,\theta)-S(y,\theta)$. By Lemma 1 this is a concave function and by definition $g(0)=0$. The result follows immediately □.

Since results for stationary dynamic programing will be used in the proof of the next two Propositions, in order to translate the non stationary problem into a stationary one we must first define a space that includes all possible realizations of the process $\{\theta_t\}$. The construction parallels the one provided in Bertsekas and Shreve (pgs.242–46) and some of the results presented there are used. Let $\Phi = \cup_{t=0}^{\infty} \Theta^t$ endowed with the topology that makes the mappings $\varphi_t:\Theta^t \to \Phi$ given by $\varphi_t(\theta^t)=\theta^t$ a homeomorphism. We will use θ^t (with variable t) to denote a typical element of Φ and λ to denote its distribution. With this topology Φ is a borel space. Note that $\mathscr{y}_t$ is a universally measurable function from $\Theta^t \to \mathcal{M}(S \times Y)$ if and only if the corresponding function $\mathscr{y}$: $\Phi \to \mathcal{M}(S \times Y)$ is universally measurable on Φ. The same holds for e_t and M_t.

Let $V_t(\theta^t,\mu_t) = \sup_{\{y_{t+\tau}\}} E_t\{\sum_{\tau=0}^{T} \beta^\tau p_{t+\tau} y_{t+\tau} \mid \theta^t\}$ on the set of feasible allocations starting from (θ^t,μ_t) with bounded value, as described in 4.1 . Note that since the technology set is linear in $M_{t+\tau}$, no increase in the value of the objective can be obtained by increasing $M_{t+\tau}$. Thus the values above can be achieved even with the restriction $M_\tau=0$ for all $\tau>t$.

We now show that this derived stochastic control problem (with $M_{t+\tau}= 0$ for all τ) satisfies the basic definitions of an optimal stochastic control model in Bertsekas and Shreve. We will refer to the following model by SM1.

i) The state space for this problem is $X=\Phi\times\mathcal{M}(S)$, which is a borel space. The controls are $(\mathcal{y},e)$ in the space $C=\mathcal{M}(S\times Y)\times\mathcal{M}(S\times Y)$ which is also a borel space, where for a measurable space Z, $\mathcal{M}(Z)$ denotes the space of positive finite measures on Z.

ii) Let $U(\theta^t,\mu)=\{(\mathcal{y},e)\in C \mid \mathcal{y}$ has first marginal μ and support on the graph of $Y(\theta^t,\cdot)$ and $e\leq \mathcal{y}\}$ We now show that the graph of this correspondence, denoted by Γ, is an analytic set. If it were not for the support restriction, this would be an immediate implication of weak convergence. Let $\varphi(s,y,\theta^t)$ be the indicator function of $\{s,y \mid y\notin Y(\theta^t,s)\}$; since $Y(\cdot)$ has measurable graph, this function is also measurable. Abusing notation, let $\tau(ds,dy,\mathcal{y})=\mathcal{y}(ds,dy)$. The function τ is trivially measurable in $\mathcal{y}$. By Lemma 2 in Appendix 5 of Dynkin and Yuskevich the function $(\theta^t,\mathcal{y})\mapsto\int\varphi(s,y;\theta^t)\tau(ds,dy,\mathcal{y})$ is borel measurable. This function has zero value for all $\mathcal{y}$ with support on $\{(s,y):y\in Y(\theta^t,s)\}$ and hence the set of pairs $(\theta^t,\mathcal{y})$ that satisfy this condition is borel measurable, so Γ is analytic (actually borel measurable).

iii) Let the disturbance space be Θ, which is a borel space and the disturbance kernel be $F(d\theta;\theta^t)$, where F is the conditional distribution of the process $\{\theta_t\}$, so it is borel measurable.

iv) The system function takes $(\mu_t,\theta^t,\theta_{t+1},\mathcal{y},e)\mapsto(\theta^{t+1},\mu_{t+1})$, where $\theta^{t+1}=(\theta^t,\theta_{t+1})$ and μ_{t+1} satisfies the feasibility condition F.4 given in section 2 (with M_t set to zero), which by assumptions (S) define a borel measurable function.

v) The return function $p(\theta^t)(\int y d\mathscr{y})$ is the composition of borel measurable functions and thus borel measurable.

Let $\Pi=(\Pi_1,\Pi_2,...)$ $\Pi_t:X\to\mathcal{P}(C)$, where $\mathcal{P}(C)$ denotes the set of probability measures on C)denote a policy for the above stochastic model, as defined in Bertsekas and Shreve (chapter 9), and denote by $V_\Pi(\theta^t,\mu)$ the value under this policy.

Define $v_t(\theta^t,s)$ as in (3.1). Given a price process $\{p_t\}$ the problem defined by (3.1) can be easily checked to satisfy the definition of an optimal control model with state space $X=\Phi\times S$ and control space Y. Let $\pi=(\pi_1,\pi_2,...)$ $\pi_t:X\to\mathcal{P}(Y)$ be a policy for this stochastic model and let $v_\pi(\theta^t,s)$ denote the value under policy π. We will denote this model SM2.

There is a correspondence between policies for SM1 and SM2. To illustrate this, consider the 'production' part of a policy Π for SM1, which is given by the universally measurable function $\mathscr{y}$: $\Phi\to\mathcal{M}(S\times Y)$. By Proposition 7.27 in Bertsekas and Shreve there exist stochastic kernels $\pi(dy;\theta^t,s)$ and $\mu(ds;\theta^t)$ such that $\pi(\cdot)$ is $\mathcal{U}(\Phi)\otimes\mathcal{B}(S)$ measurable and $\mu(ds;\theta^t)$ is $\mathscr{U}(\Phi)$ measurable, where $\mathscr{U}$ and $\mathscr{B}$ denote respectively, the universal and borel σ–algebras, that satisfy:

$$\mathscr{y}(A\times B;\theta^t) = \int_A \pi(B;\theta^t,s)\mu(ds;\theta^t) \quad \text{for all sets } A\in\mathscr{B}(S) \text{ and } B\in\mathscr{B}(Y).$$

Note that this formula implies $\mu_t(ds)=\mu(ds;\theta^t)$ and that for almost all (θ^t,s) $\pi(Y(\theta^t,s);s,\theta)=1$, since $Y(\theta^t,s)$ is the section at s of Gr $Y(\theta^t,\cdot)$. By appropriately altering π on this measure zero set[21] we obtain a policy for SM2 that corresponds to $\mathscr{y}$.

In fact, a deterministic model (D) can be defined following the construct in Bertsekas and Shreve, section 9.2, that corresponds to both the SM1 and SM2 models. We refer the reader to the reference. Let Π' denote a policy for (D).

With these elements in mind we will prove Proposition 3.

[21]This can be done since there exists a universally measurable selection from the set of feasible input output choices.

Proof of Proposition 3. Since model (D) corresponds to both SM1 and SM2, and by (B3') and (SES) or (SES') the return function is bounded above, Proposition 9.5 in Bertsekas and Shreve imply that

(*) $$V(\theta^t,\mu_t)=\int v(\theta^t,s)\mu_t(ds).$$

where $V(\theta^t,\mu_t)=\sup_{\Pi} V_{\Pi}(\theta^t,\mu_t)$ and $v(\theta^t,s)$ is defined in a similar way. This equation provides the key connection between the single firm equilibrium and the Industry equilibrium.

If Π is an optimal policy for SM1, there is a corresponding policy π' for (D) and thus a policy π corresponding to SM2, all of which are optimal (Corollary 9.5.1). By construction, π gives the conditional distributions for input output choice and exit of Π, and since it is optimal, it must have support on profit maximizing choices. Hence a single firm equilibrium satisfies (ii) and (iii) of the definition of Industry equilibrium. Also note that by (*)

$$V(\theta^t,\mu_t+M\nu) = V(\theta^t,\mu_t)+MV(\theta^t,\nu) = V(\theta^t,\mu_t)+M\int v(\theta^t,s)\nu(ds)$$

so condition (iv) is verified. Thus any single firm equilibrium is an Industry equilibrium. Conversely, given an Industry equilibrium with optimal policy π for SM2, the corresponding policy Π for SM1 is optimal. It is then easy to check that the Industry equilibrium is also a Single firm equilibrium □.

We now establish the existence of a solution to the optimal program . This problem is identical to the single firm problem described above but with a different objective and without the restriction of $M_t=0$. It is easy to check that with these modification (i)–(iv) satisfy the definition of a stochastic control model. The last condition is also satisfied replacing (v) by:

(v′) The return function $S(\theta^t,(\int y\, d\mathcal{Y}))$ which applying Lemma 1, is the composition of borel measurable functions and thus borel measurable. To simplify notation we refer to this function as $S(\theta^t,\mathcal{Y})$.

To prove the existence of an optimal solution we show that the set

$$U_k(\theta^t,\mu_t,\lambda) = \left\{(M_t,\mathscr{y}_t,e_t)\in U(\theta^t,\mu_t) \mid S(\theta^t,\mathscr{y}) + \beta\int V^k(d\theta^{t+1},\mu_{t+1}\mid\theta^t) \geq \lambda\right\}$$

where $\mu_{t+1}, \mathscr{y}_t$ and e_t satisfy F.4 in the definition of feasiblity, is compact in C, where V^k corresponds to the k–th iterate on the optimality equation and $V^0=0$. By Proposition 9.17 in Bertsekas and Shreve, this implies the existence of an optimal nonrandomized stationary policy.

We first establish that the optimality equation maps bounded measurable functions that are upper semi–continuous in μ into the same class. By (B1)–(B3) the return function $S(\theta^t,\cdot)$ is bounded above. We will prove that it is upper semicontinuous.

We first consider case (SES), i.e. when Y is bounded above. Given that $S(\theta^t,\cdot)$ is continuous, it suffices to show that whenever $\mathscr{y}^n \to \mathscr{y}$, $\limsup_n y^n \leq y$ where $y^n=\int y\mathscr{y}^n(ds,dy)$ and y is similarly defined. Let $c_m\in R^k_-$ be a sequence decreasing to $-\infty$ and let χ_m denote the indicator function of $[y\geq c_m]$. Choose m so that $\int y\, \mathscr{y}(ds,dy)+\epsilon > \int \chi_m y \mathscr{y}(ds,dy)$. Since $\chi_m y$ is an upper semicontinuous function of y and by (SES) it is bounded on the domain of integration, $\mathscr{y}^n \to \mathscr{y}$ implies

$$\limsup_n \int y\mathscr{y}^n(ds,dy) \leq \limsup_n \int \chi_m y\, \mathscr{y}^n(ds,dy) \leq \int \chi_m y\mathscr{y}(ds,dy) \leq \int y\mathscr{y}(ds,dy)+\epsilon.$$

Since ϵ can be chosen arbitrarily small, the result follows.

We now consider the alternative assumption (SES'). By (B1) there exists s^* such that $S_i(\theta^t,\cdot)\leq s^*/k$, where k is the number of goods (inputs and outputs) from the definition of an Industry. We now show that for $a\leq s^*$ the set $\{y \mid y=\int y\mathscr{y}(ds,dy)$ for some $\mathscr{y}$ with support on gr $Y(\theta^t,\cdot)$ and $S(\theta^t,y)\geq a\}$ is closed. Let $y^n\to y$ be a sequence satisfying the condition. Let $\{\mathscr{y}^n\}$ be the associated sequence of measures. This sequence is tight[22] so, if necessary along a subsequence, it converges to some measure $\mathscr{y}$, which also has support on

[22]Tightness results from the fact that $S(\theta^t,y^n)\geq a$ implies that the sequence y^n is norm bounded. It is easy to check that this implies that $\mathscr{y}^n$ is a tight sequence of measures.

gr $Y(\theta^t,\cdot)$. For $\epsilon>0$ choose $c>0$ so that $\|y\|\geq c$ implies $\|y\|/|y_j| < \frac{p_j \epsilon}{a+s^*}$ and let χ_c be the indicator function of the set $[y|y_i<c]$. As before $\chi_c y$ is an upper semi continuous function on R^k. Note that $\int_{\|y\|\geq c} \|y\| \mathcal{Y}^n(ds,dy) < \epsilon$, for otherwise $S_j(\theta^t,y^n) \leq p_j \int y_j \mathcal{Y}(ds,dy) < -(a+s^*)$. In consequence

$$\text{limsup} \int \mathcal{Y}^n(ds,dy) - \epsilon \leq \text{limsup} \int \chi_c \mathcal{Y}^n(ds,dy) \leq \int \chi_c \mathcal{Y}(ds,dy) \leq \int \mathcal{Y}(ds,dy),$$

so $S(\theta^t,y)\geq a$ and $S(\theta^t,\cdot)$ is upper semi continuous.

Since (B1) implies that V^k is bounded above by an integrable function, a direct application of Fatou's lemma implies that $\int V^k(d\theta^{t+1},\mu_{t+1}|\theta^t)$ is upper semi continuous in μ_{t+1} whenever $V^k(\theta^{t+1},\mu^{t+1})$ is. Finally note that given θ^t , μ_{t+1} is linear in $\mathcal{Y}_t$ and e_t so for all θ^t the objective of the functional equation (and also in the definition of U_k) is upper semicontinuous (an bounded) in the controls. Furthermore, by (B2) the definition of U_k implies that the input output vectors associated to $\mathcal{Y}_t$ must be normed bounded, so $\mathcal{Y}_t$ must belongs to a tight set and that, by condition (T3), $M_t\leq M$ for some finite M. This implies that the set of $(M_t, \mathcal{Y}_t, e_t)$ that make the objective greater than or equal to λ is compact.

To establish that $V^{k+1}(\theta^t,\cdot)$ is upper semicontinuous, it suffices to show that $U(\theta^t,\cdot)$ has closed graph. This also completes the proof of existence since it implies that $U(\theta^t,\mu_t)$ is a closed set. Without the support restriction using the definition of weak convergence of measures it is easy to check that the graph of $U(\theta^t,\cdot)$ is closed. To verify that the support restriction, note that since the graph of $Y(\theta^t,.)$, which we will denote by Γ, is closed, $\mathcal{Y}^n \to \mathcal{Y}$ implies $\mathcal{Y}(\Gamma) \geq \text{limsup}\ \mathcal{Y}^n(\Gamma)$ and obviously $\mathcal{Y}(S\times Y) = \lim \mathcal{Y}^n(S\times Y)$. Thus $\mathcal{Y}^n(\Gamma) = \mathcal{Y}^n(S\times Y)$ implies that $\mathcal{Y}$ has also support on Γ, and the proof is complete □.

<u>Lemma 3</u>. Given a compact metric space S and a complete, separable metric space T consider their product S×T with the product σ–algebra of their respective borel σ–algebras. For each measure μ on S let $\gamma(\mu)$={measures ν on S×T with marginal μ on S}. Then γ is a lower hemi–continuous correspondence.

Proof. Given $\mu_n \to \mu$ and $\nu \epsilon \gamma(\mu)$ we must construct a sequence $\nu_n \epsilon \gamma(\mu_n)$ with $\nu_n \to \nu$. Intuitively, for each μ_n we want to 'distort' ν minimally so that its marginal coincides with μ_n instead of μ. Since $\mu_n \to \mu$ we may expect that this distortion disappears in the limit.

Since the spaces S and T are separable metric spaces, they have countable bases B_s and B_t with A μ–continuous for all $A \epsilon B_s$ and B $\nu(S,.)$–continuous for all $B \epsilon B_t$. Let $S_m = \{A_{mj}:j=1,..,m\}$ and $T_m = \{B_{mk}:k=1,...,m\}$ be the exhaustive sequence of subdivisions of S and T constructed from B_s and B_t as described in Lemma 6. Thus $S_{m+1}(T_{m+1})$ is a refinement of $S_m(T_m)$, each set $A \epsilon B_s$ ($B \epsilon B_t$) is the finite union of sets in $\bigcup_{s=1}^{\infty} S_m(\bigcup_{s=1}^{\infty} T_m)$ and the sets A_{mj} and B_{mj} are μ–continuous and $\nu(S,.)$–continuous, respectively.

Suppose $A \epsilon B_s$ and $B \epsilon B_t$. Then we can find partitions S_m and T_m such that A is the union of sets in S_m and B of elements in T_m. Thus $A \times B$ is the finite union of rectangles in $S_m \times T_m \equiv \{A_{mj} \times B_{mk}\}_{j,k=1}^{m}$. Thus the collection of sets $\{S_m \times T_m\}_{m=1}^{\infty}$ is a basis for the product topology in $S \times T$ and hence is convergence determining (see Theorem 2.2 in Billingsley).

Let $\alpha_{mjk} = \nu(A_{mj} \times B_{mk}) / \nu(A_{mj} \times T)$ whenever $\mu(A_{mj}) \neq 0$ and 0 otherwise. For each (m,k) choose $t_{mk} \epsilon B_{mk}$. Define the measure ν_{nm} on $S \times T$ by setting

$$\nu_{nm}(C) = \sum_{j=1}^{m} \sum_{k=1}^{m} \alpha_{mjk} \mu_n(\pi_s(C \cap A_{mj} \times \{t_{mk}\}))$$

for each C in $\mathcal{S} \otimes \mathcal{T}$, where π_s is the projection function on S. Define ν_m similarly with measure μ instead of μ_n. The set functions ν_{nm} and ν_m are clearly positive measures since μ_n and μ are. For any A in $\mathcal{S}$, $\pi_s(A \times T \cap A_{mj} \times \{t_{mk}\}) = A \cap A_{mj}$. In consequence,

$$\nu_{nm}(A \times T) = \sum_{j=1}^{m} \sum_{k=1}^{m} \alpha_{mjk}\, \mu_n(A \cap A_{mj}) = \mu_n(A)$$

so μ_n is the marginal of ν_{nm} and by a similar argument μ is the marginal of ν_m.

We will now show that $\nu_{nm} \to \nu_m$ and that ν_m converges to ν. Let A be a μ–continuity set and B a $\nu_m(S,.)$–continuity set. Suppose $A \times B \cap A_{mj} \times \{t_{mk}\} \neq \phi$. Then $A \times B \cap A_{mj} \times \{t_{mk}\} = A \cap A_{mj} \times \{t_{mk}\}$. Since A and A_{mj} are μ–continuity sets, so is $A \cap A_{mj}$. Hence $\mu_n(A \cap A_{mj}) \to \mu(A \cap A_{mj})$ and thus $\nu_{nm}(A \times B) \to \nu_m(A \times B)$, so $\nu_{nm} \to \nu_m$ (Theorem 3.1 in Billingsley).

Note that $\nu_m(A_{mj}\times B_{mk})=\alpha_{mjk}\mu(A_{mj})=[\nu(A_{mj}\times B_{mk})/\nu(A_{mj}\times T)]\mu(A_{mj})=\nu(A_{mj}\times B_{mk})$.

By additivity of ν it is also true that $\nu_{m'}(A_{mj}\times B_{mk})=\nu(A_{mj}\times B_{mk})$ for all $m'\geq m$. Consequently $\lim_{m'\to\infty}\nu_{m'}(A_{mj}\times B_{mk})\to\nu(A_{mj}\times B_{mk})$ for all sets A_{mj} and B_{mk}. Since this class of sets is convergence determining, $\nu_m\to\nu$.

Our final step is to extract a sequence ν_n converging to ν with marginals μ_n. Let $N(m)$ be such that if $n\geq N(m)$ then $\rho(\nu_{nm},\nu_m)<\frac{1}{m}$. Without loss of generality choose $N(m)$ to be increasing in m. For $n=1,\ldots,N(1)$ let $\nu_n=\nu_{n1}$. For $N(m)\leq n<N(m+1)$ let $\nu_m=\nu_{nm}$. Let $\epsilon>0$ and choose M so that $\rho(\nu_m,\nu)<\epsilon/2$ for all $m\geq M$ and $\frac{1}{M}<\epsilon/2$. Then for all $n\geq N(m)$ we have

$$\rho(\nu_n,\nu)\leq\rho(\nu_n,\nu_m)+\rho(\nu_m,\nu)<\frac{1}{m}+\epsilon/2\leq\epsilon \text{ for some } m\geq M \ \square.$$

<u>Lemma 4</u>. Let $\gamma:S\twoheadrightarrow T$ be a continuous correspondence with closed and convex values from a paracompact space S to a Banach space T. There exists a continuous function $h:S\times T\to T$ such that:

i) $h(s,t)\in\gamma(s)$)

and ii) $h(s,t)=t$ for $t\in\gamma(s)$.

<u>Proof</u>. Define the correspondence $\Gamma:S\times T\to T$ by letting $\Gamma(s,t)=t$ for all $(s,t)\in gr(\gamma)$ and $\Gamma(s,t)=\gamma(s)$ otherwise. Γ is lower hemi–continuous, with closed and convex values. By Michael's selection theorem Γ has a continuous selection $h:S\times T\to T$. This function satisfies conditions i) and ii) $\square$.

Let Γ be a correspondence defined by

$\Gamma(x)=\{y\in\mathbf{M}:0\leq y\leq x\}$

where $x\in\mathbf{M}$ and $\mathbf{M}$ is the space of positive measures on compact metric space S with total mass uniformly bounded..

Lemma 5. Γ is lower hemi–continuous.

Proof. Suppose $x_n \to x$ and $0 \leq y \leq x$. Without loss of generality we may consider $x \neq 0$ We must find a sequence y_n such that $y_n \to y$ and $0 \leq y_n \leq x_n$. Note that if the space were discrete then letting $y_n(j)=\min\{x_n(j),y(j)\}$ the sequence y_n would satisfy the above. For an arbitrary space we will construct an exhaustive sequence of subdivisions to prove via an approximation the existence of such a sequence y_n.

Let $\{P_m\}$ denote an exhaustive sequence of subdivisions and set $P = \bigcup_{m=1}^{\infty} P_m$. We will denote a typical element of subdivision P_m by A_{mk}. By Lemma 6 we can assume that $\{P_m\}$ satisfies:

i) Every set $A_{mk} \epsilon P$ is an x–continuity set.

ii) P is a convergence determining class.

Define $y_m \epsilon \mathbf{M}$ by letting $y_m(A)=[y(A_{mk})/x(A_{mk})]x(A)$ whenever $A \subset A_{mk}$. This is the measure x scaled down by a factor that depends on the y measure of the sets of the partition relative to the x measure. For each pair (n,m) we will define a measure y_{mn} by following a similar scaling procedure. For this purpose, let

$\alpha_{mkn}=\min\{y(A_{mk}),x_n(A_{mk})\}/x_n(A_{mk})$ if $x_n(A_{mk})>0$ and set it to zero otherwise.
Let $y_{mn}(A)=\alpha_{mnk}x_n(A)$ if $A \subset A_{mk}$. Clearly $0 \leq y_{mn} \leq x_n$.

We now show that $\lim_{n\to\infty} y_{mn}=y_m$. We will show that $y_{mn}(A) \to y_m(A)$ for all sets A such that $y_m(\partial A)=0$. Without loss of generality we can restrict attention to a set A contained in some A_{mk}. Notice that for A to be a continuity set of y_m it must be the case that either $y(A_{mk})=0$ or that $x(\partial A)=0$. The first case is trivial. In the second case $x_n(\partial A) \to x(\partial A)$. But since A_{mk} is a continuity set of x, $x_n(A_{mk})$ also converges to $x(A_{mk})$ so α_{mkn} converges to $y(A_{mk})/x(A_{mk})$. In consequence $y_{mn} \to y_m$ as desired.

We now show that $y_m \to y$. Notice that any set $A_{mk} \epsilon P$ is the finite union of sets in $P_{m'}$ for $m' \geq m$ so by construction of y_m, $y_{m'}(A_{mk})=y(A_{mk})=y(A_{mk})$. Since P was chosen to be convergence determining we can conclude that $y_m \to y$.

We will now use a 'diagonal' type argument to generate the sequence y_n. Let N(m) be such that for $n \geq N(m)$ $\rho(y_{mn},y_m)<\frac{1}{m}$. Without loss of generality take N(m) to be strictly increasing. For $n<N(2)$ choose $y_n=y_{1m}$ and for $N(m)\leq n<N(m+1)$ choose $y_n=y_{mn}$. Let $\epsilon>0$. Let m be such that $\rho(y_m,y)<\epsilon/2$ and $\frac{1}{m}<\epsilon/2$. Choose $n\geq N(m)$. Then $y_n=y_{nm'}$ for $m'\geq m$ and $n\geq N(m')$. In consequence $\rho(y_n,y)\leq\rho(y_{nm'},y_{m'})+\rho(y_{m'},y)<\epsilon$. Hence $y_n\to y$ and $0\leq y_n\leq x_n$ so Γ is lower hemi–continuous □.

Lemma 6. Suppose S is complete separable metric space and $x\in\mathcal{M}(S)$. Then there exists an exhaustive sequence of subdivisions $\{P_m\}$ of S [23] such that

i) All sets in each P_m are x–continuity sets.

ii) $P = \bigcup_{m=1}^{\infty} P_m$ is a convergence determining class.

Proof. Since S is separable metric, the collection of open sets that are x–continuous is a base for the topology of S and the base can be chosen to be countable. Define the exhaustive sequence of partitions P as done in Gihman and Skorohod. By construction, these sets are x–continuity sets. P is closed under the formation of finite intersections and each open set in S is a finite or countable collection of elements of P. By Theorem 2.2 in Billingsley P is a convergence determining class, i.e. if $\mu_n(A)\to\mu(A)$ for all $A\in P$ we may conclude that $\mu_n\to\mu$ □.

[23]For definitions and results see Gihman and Skorohod, especially pg. 63 and Lemma 5.

References

Bertsekas and Shreve. *Stochastic Optimal Control: The Discrete Time Case.* Academic Press, 1978.

Davis, S. and J. Haltiwanger. "Gross Job Creation, Gross Job Destruction and Employment Reallocation," Unpublished manuscript, January 1989.

Duffie D., J. Ganakoplos, A. Mas–Colell, and A. McLennan "Stationary Markov Equilibria," Research Paper No. 1079, Graduate School of Business, Stanford University.

Dunne T., M. Roberts and L. Samuelson (1987a). "Patterns of Firm Entry and Exit in US Manufacturing Industries," Department of Economics, The Pennsylvania State University, 1987.

________________ (1987b). "Plant Failure and Employment Growth in the US Manufacturing Sector," Department of Economics, The Pennsylvania State University, 1987.

Ericson R. and A. Pakes (1989) "An Alternative Theory of Firm and Industry Dynamics, manuscript.

Evans, D.S. (1987a) "The Relationship Between Firm Growth, Size and Age: Estimates for 100 Manufacturing Industries", Journal of Industrial Economics, June 1987.

Evans, D.S. (1987b) "Test of Alternative Theories of Firm Growth," Journal of Political Economy Vol. 95, pp. 657–674.

Hildenbrand, W. (1975) "Distribution of Agents' Characteristics," Journal of Mathematical Economics 2(1975).

Hopenhayn, H. (1989) "A Dynamic Stochastic Model of Entry and Exit to an Industry," Research Paper # 1057, Graduate School of Business, Stanford University.

Hopenhayn, H. and R. Rogerson, "Job Turnover and Policy Evaluation in a Model of Industry Equilibrium," mimeo.

Jovanovic, B. (1982) "Selection and the Evolution of Industry." Econometrica, vol. 50 N^{o}3 (May 1982).

Lucas, R and E.C. Prescott (1971) "Investment Under Uncertainty," Econometrica 39(1971).

Novshek, W. and H. Sonnenschein, "Cournot and Walras Equilibrium," JET 19:223–226, 1978.

Pakes A. and R. Ericson. "Empirical Implications of Alternative Models of Firm Dynamics," Working Paper No 2893, National Bureau of Economic Research, March 1989.