COMPUTING MARKOV PERFECT NASH EQUILIBRIA: NUMERICAL IMPLICATIONS OF A DYNAMIC DIFFERENTIATED PRODUCT MODEL

Ariel Pakes*   Paul McGuire*

Yale University   Yale University

and National Bureau of Economic Research   and Economic Growth Center

Abstract

This paper provides an algorithm for computing Markov Perfect Nash Equilibria (Maskin and Tirole, 1988a and b) for dynamic models that allow for heterogeneity among firms and idiosyncratic (or firm specific) sources of uncertainty. It has two purposes. To illustrate the ability of such models to reproduce important aspects of reality, and to provide a tool which, given appropriate parameter estimates, can be used for both descriptive and policy analysis in a setting which allows firms to differ from one another in ways that are consistent with the information in firm level data sets.

We illustrate by computing the policy functions, and simulating the industry structures, generated by a class of dynamic differentiated product models in which the idiosyncratic uncertainty is due to both the random outcomes of each firm's research process, and to an autonomous aggregate demand process. The illustration focuses on comparing the effects of different regulatory and behavioral assumptions on market structure and on welfare for one particular set of parameter values. The results here are of some independent interest and can be read without delving into the technical detail of the computational algorithm.

The last part of the paper begins with an explicit consideration of the computational burden of the algorithm, and then introduces approximation techniques designed to make computation easier. The purpose of this section is to enable us to compute equilibria for industries in which a large number of firms are typically active. Its major result is analytic. We show that if the value function of a given firm is exchangeable in the state vectors of its competitors, then the number of polynomial coefficients one needs for a given order of a polynomial approximation to that function is both independent of the number of firms active in the market, and a relatively small number. This enables us to use the approximation technique to reduce the computational burden of the algorithm dramatically.

This paper is part of an ongoing project designed to help structure the empirical analysis of data on firms and industries. It has two purposes; to illustrate the ability of models that allow for heterogeneity among firms and idiosyncratic (or firm specific) sources of uncertainty to reproduce important aspects of reality, and to provide a computational framework that allows one to compute the equilibrium implications of these models. As we illustrate below, provided one has estimates of the parameters needed for the computations, this algorithm can be used for analyzing both the descriptive and welfare implications of alternative policy and environmental changes. For a discussion of some of the issues that would arise in estimating these parameters, see Pakes (forthcoming) and the literature cited their.[2]

One of the more striking features of data on plants or firms is the degree of heterogeneity among firms in the same "industry" in both the levels, and the movements over time, in the variables we typically want to analyze (shares in industry output, investment, productivity, etc.). Most often these differences have a direct impact on the way we analyze the issues of interest. That this is so even if we are only after the "aggregate" impacts of a policy or an environmental change (say the effect of a change in input prices on industry output, or the effect of an investment tax credit on productivity) is a result of the simple fact that firms' responses in realistic models of behavior are typically *different nonlinear* functions of the changing variable. As a result the industry's aggregate response, which is the sum of the individual responses, depends on higher order moments of the distribution of individual responses.

In markets with little entry and exit the differences in response patterns are largely a result of differences in the past choices and the past successes of incumbents, and

---

[2]Two points on estimation are probably worth noting here. First, estimation is separate from computation in the sense that one can estimate the relevant parameters without computing their equilibrium implications. Second, estimates of most of the needed primitive parameters can be obtained from the cross sectional variation in panel data sets, and do not require a long time series for their consistency properties.

nonlinearities are induced by equilibrating mechanisms which typically imply that losses to one set of firms are partially offset by gains to others. In addition entry and exit are examples of an extreme nonlinearity in response patterns, and the extent of both these processes typically depends directly on the change we want to analyze. So any analysis of the aggregate impacts of a policy or environmental change on a market requires not only the distribution of responses of the incumbents in the given market, but also an analysis of the entry and exit that the policy or environmental change induces. Interestingly the importance of explicitly accounting for heterogeneity in response patterns when analyzing the aggregate impacts of changes has been a theme of recurrent emphasis in the (recent) empirical literature that uses panel data on firms or plants to study an industry's reaction to a policy or an environmental changes (for two examples see Thomas',1990, analysis of the impact of changes in FDA regulations on the rate of innovation in the pharmaceutical industry, and Olley and Pakes',1991, study of the impact of deregulation on aggregate productivity growth in the telecommunication equipment industry).

Moreover, we are often specifically interested in the relationship of policy, or environmental, variables to the more detailed structure of the distribution of agents characteristics, and in how the equilibrium distribution of those characteristics is likely to react to the policy or environmental change. Obvious examples where more detailed knowledge of the determinants of the equilibrium distribution is of overriding importance are easy to come by in almost all aspects of economics. The analysis of the link between default probabilities and the market for finance capital, and of the effects of the various regulatory changes on market structure, are examples that occur repeatedly in the finance and industrial organization literatures; and we focus on the effects of regulatory changes in the numerical analysis below. More recently, the finding that almost all of the variance in gross job creation and gross job destruction is within time—period, within—industry, variance (see Davis and Haltwinger,1989), makes any analysis of the causes or the effects of job turnover in labor markets highly dependent on the detailed characteristics of the

equilibrium from dynamic heterogeneous agent models (for an analysis of this dependence in one particular setting see Hoppenhayn and Rogerson,1990). We note that once our models acknowledge the fact that agents do differ, and grant that their actions may impact on one another, then the computation of either the aggregate response, or of the distributional characteristics, needed to analyze any of these problems can become quite demanding.[3]

There is more than one source of idiosyncratic uncertainty that firms react to, and different sources are likely to be more relevant for analyzing behavior in different industries. We focus here on a differentiated product version of the Ericson—Pakes (1990) model of industry dynamics. That model focuses on the uncertainty generated from the random outcomes of research and exploration processes, and so should be primarily applied to industries where these processes seem to be important. [4]

The Ericson—Pakes model has firms invest to explore, and if warranted, develop, profit opportunities (improved goods or techniques of production). The outcomes of the investment process are uncertain. Conditional on an industry structure, positive outcomes lead the firm to states where the good or service can be marketed more profitably. If the outcomes generate lesser increments than those of its competitors (both inside and outside the industry) the firm's profits deteriorate, and may lead to a situation in which it is optimal to abandon the whole undertaking (this endogenizes exit behavior, and provides

---

[3]The Hoppenhayn and Rogerson paper is also one of the few that worries about computation in heterogeneous agent models. They assume that all agents are zero measure and all sources of uncertainty are idiosyncratic, show that under their conditions the industry structure converges to a fixed s* (and stays there) and then provide a simple way of computing s*. Judd (1990) has computed Markov perfect equilibria for two agent models with no entry and exit, and Hansen and Sargent (1990) provide a computational algorithm for a class of heterogeneous agent models that allow for linear decision rules and equilibria (they assume quadratic preferences, linear technologies and information sets, no discrete choices, and that continuous choices are always interior).

[4]It is, at least analytically, quite easy to incorporate exogenous price uncertainty provided the alternative possible price realizations do not reorder the relative efficiencies of the firms in the industry, as in Dixit, 1989. On the other hand, if relative efficiencies were effected by price realizations, as in Lambson (forthcoming), we would require a more substantive remodelling effort.

one way of accounting for selection in the nature of the evolutionary process).

A firm's supply to the spot market for current output, and its current profits, depend on its own level of development, a counting measure which provides the level of development of other firms currently active in the industry, and the level of development of the alternatives outside the industry. The level of development of the outside alternative evolves exogenously. Entry and investment decisions (which determine the levels of development of the actors in the industry) are made to maximize the expected discounted value of future net cash flow conditional on the current information set. That information set includes a distribution for the counting measure of possible industry structures in future years conditional on the current structure. The equilibrium notion insists that this distribution is in fact consistent with optimal investment behavior by all incumbents and potential entrants. Since we restrict strategies to be measurable functions of the set of variables which determine either current production costs or current demand conditions (to use the terminology of Tirole, 1989, to "payoff relevant" state variables), the equilibrium we obtain is Markov—Perfect Nash in investment strategies in the sense of Maskin and Tirole (1987,1988a,1988b).

At the heart of this equilibrium is a stochastic process which generates the industry structures that the model emits. That process is ergodic, but its more detailed characteristics depend on the precise values of the model's parameters. So we introduce an algorithm for computing the equilibria, and use it to analyze special cases of the model. Here we begin by simply simulating out the pattern of industry structures that are likely to materialize for different sets of parameters. This allows us to illustrate the many aspects of reality that can be captured by models which allow for idiosyncratic uncertainties in a dynamic market setting.

We then move on to a numerical analysis of the effects of different institutional arrangements on market structure and on welfare for one particular set of technological parameter values. The analysis compares five different institutional structures to one

another, and to the outcomes of the actions of a social planner. Those institutional structures are; the free entry Markov Perfect equilibrium, an initial temporary monopoly followed by free entry, free entry with an institutional constraint which limits the market share of the largest firm to be below some fixed fraction of the overall market, free entry with institutionally induced sunk entry costs (a license must be obtained to enter the market), and a "perfect" cartel. The results from these experiments provide a comparison of the (sometimes complex but always intuitive) impacts of the institutional arrangements analyzed, and illustrate the potential usefulness of the computational techniques for policy and descriptive analysis.

The algorithm we suggest for computing the Markov Perfect Nash equilibria is analytically simple, but may be computationally demanding, particularly if the industry we are describing tends to have a large number of firms active in a given period. The last part of the paper begins with an explicit consideration of the computational burden of the algorithm, and then introduces approximation techniques designed to make computation easier. We focus on techniques based on polynomial approximations to the value function.

Define a grid point to be a combination of the value of a firm's own state variable and a list of the state variables of the firm's active competitors. The approximation techniques are based on fitting a polynomial to the value of being active at a small number of grid points, and then using the polynomial coefficients obtained from this fit to predict the value function at other points as needed.

The major result of this section is analytic. It shows that provided the value function of a given agent is symmetric (more precisely exchangeable) in the state vectors of its competitors, then the number of polynomial coefficients one needs to determine for a given order of approximation is independent of the number of agents active in the market. This result implies that the number of grid points we need to evaluate at each iteration of the fixed point calculation does not depend on the number of firms active in the industry. Unfortunately the computational burden at each grid point still depends on the number of

active firms. We do, however, show how one can adapt a technique due to Kortum (forthcoming) to reduce the computational burden of evaluating each grid point from being (a high order) polynomial in the number of firms that are active at that grid point, to being linear in the number of active firms. Hopefully this combination of results will allow us to use the algorithm to analyze markets with large numbers of firms. Whether or not it does depends on the quality of the resulting approximations, and this may well vary from problem to problem. We conclude with a brief summary of some early results on the fit of the approximation for our problem.

The paper begins by providing a brief description of each of the primitives of the model, starting with the profit function, then turning to the other primitives determining incumbent behavior, and finally to those determining the behavior of entrants. For each primitive we outline the general characteristics it need possess, and then go over a detailed example that is used in the actual computations. We then give a verbal characterization of optimal policies and the nature of equilibrium (for more detail see Ericson and Pakes,1989). The second section of the paper introduces the computational algorithm, the third provides the numerical results, and the fourth considers its computational burden and the approximation techniques. A reader should be able to read sections I and III of this paper without paying any attention to sections II and IV.

I. The Model.

The state variables determining the firm's perception of its opportunities are

$$(\omega,s)\epsilon \ \Omega \ X \ S \subseteq \ \mathscr{Z} \ X \ \mathscr{Z}_+^{\mathscr{Z}},$$

where $\mathscr{Z}$ is notation for the integers, so ω is an integer–valued index of the firm's own efficiency, and the vector s is a counting measure providing the number of firms at each possible efficiency level.

Thus

$$\pi(\cdot,\cdot):\Omega X S \rightarrow \mathbb{R},$$

provides the "reduced" form of the current profit function. In the general case we need only that $\pi(\omega,s)$ is: increasing in ω for all s; that there exists a complete preorder on S, say $\succeq$, s.t. if $s_1 \succeq s_2$, then $\pi[\omega,s_1] \geq \pi[\omega,s_2]$ for all ω; and that sup $\pi[\omega,s] \leq \pi^*$ and for each ω, $\pi[\omega,s] \leq (1-\beta)\Phi$ for sufficiently large s.

Here β is the discount rate, and Φ is the scrap or exit value (the value of the firm and its entrepreneur in its best alternative use) so $(1-\beta)\Phi$ is the per period return on the firm's transferable assets. Thus the last condition states that no matter the firm's ω, there is a level of competition that is sufficiently stiff to make the returns the firm earns by operating in the industry less than the returns that would be earned were the firm's resources invested in its best alternative activity.

The special case we actually compute is a differentiated product model (for more detail on the specification of such models see Berry,1991, Berry Levinsohn and Pakes,1991, and the literature cited in these articles). Good "0" is the outside good, and goods 1,...,N are the goods produced by the firms competing in the industry. Each consumer purchases at most one good from the industry. The utility consumer "i" derives from purchasing and consuming good "j" is given by

$$(1) \quad U(i,j) = v_j - p^*_j + \epsilon(i,j),$$

where v_j is the quality or efficiency index, and p^*_j is the price, of the good, and i=1,...,M. Consumer "i" chooses good "j" if and only if he prefers it over all the alternatives, that is if for q=0,1,...,N,

$$(2) \quad \epsilon(i,j) - \epsilon(i,q) \geq [v_q - v_j] + [p^*_j - p^*_q]$$
$$= [v_q - v_0] - [v_j - v_0] + [p^*_j - p^*_0] - [p^*_q - p^*_0]$$
$$= g[\omega_q] - g[\omega_j] + p_j - p_q,$$

where $g(\cdot)$ is increasing and concave, $\omega_q = g^{-1}[v_q - v_0]$, and $p_q = p^*_q - p^*_0$. We will assume $g(\cdot)$ to have an upper bound, and this will generate the needed upper bound to the profit function.[5]

Note that (2) implies that consumer choices are determined entirely by the quality and prices of the goods marketed in this industry relative to the quality and price of the outside alternative (only relative qualities and prices count, and we have chosen the outside alternative as our "numeraire"). Therefore when we refer to an increase in the quality of the product marketed by the firm we will mean an increase relative to the outside alternative, and we will say quality decreases only when the improvements to the firm's own product are not as great as the improvements in the outside alternative. Note also that movements in v_0 will cause synchronized movements in the relative efficiencies, in the ω's, of all firms in the industry, which in turn will generate a *positive* correlation in their profits. Since movements in the v_j will tend to generate negative correlations in the profits of the firm's in an industry, dealing explicitly with the outside alternative allows us to rationalize the positive correlations in the profits of firms within an industry that we often observe in the data.

Let the set $C[\omega_j; p, s]$, where s is the counting measure providing the number of firms at each ω, be the set of ϵ's that satisfy the set of inequalities in (2), and hence induce the choice of good j. Then, letting $G(\cdot)$ provide the distribution of ϵ and assuming it to be

---

[5] g generates decreasing marginal utility to increments in the relative quality of the goods marketed in this industry, and this, in turn, generates an upper bound to profits. A more detailed model of consumers would explicitly incorporate a distribution of income constraints and decreasing marginal utility of income (see, for eg. Berry, Levinsohn and Pakes, 1991), and this would have similar effects on demand patterns (though it would also be computationally more burdensome).

multivariate extreme value, the probability that a random consumer will choose good "j" is

$$(3) \quad \sigma[\omega_j;p,s]=\int_{\epsilon \subseteq C[\omega_j;p,s]} dG(\epsilon)=\exp[g(\omega_j)-p_j]/\{1+\Sigma\exp[g(\omega_q)-p_q]\}.$$

If there are N firms in the market, no fixed costs of production and constant marginal costs equal to mc, then it can be shown that if firms choose prices to maximize profits, $(p-mc)M\sigma_j(w;p,s)$, a unique Nash equilibrium exists (Caplin and Nalebuff,1991) and satisfies the vector of first order conditions

$$(4) \quad -[p_j-mc]\sigma_j[1-\sigma_j] + \sigma_j = 0$$

for j=1,...,N. Profits are given by

$$(5) \quad \pi[\omega_j,s]=\{p[\omega_j,s]-mc\}M\sigma[\omega_j;s],$$

where it is understood that the price and share vectors are calculated from the spot market equilibrium conditions (3 and 4).

The distribution of ω_{t+1} conditional on ω_t depends on the amount the firm is willing to invest in developing its product. We let the family of distributions for the increment in ω, i.e. for $\omega_{t+1} - \omega_t = \tau_{t+1}$, conditional on different values of x_t (investment), be

$$\mathbb{P} = \{ P(\cdot|x), \ x\epsilon\mathbb{R}_+ \},$$

which we assume to have finite support.

This family is built as a difference, or a convolution, of two random variables. The

first, say v_1, is the increment in efficiency the firm gets from its own research process, and is stochastically increasing in x (in the first order dominance sense). The second, say v, is an exogenous random variable which represents the force of the competition from outside of the industry (the efficiency of the outside alternative in the example above). Note that the possibility of advances by outside competitors imply both that; $P(\cdot)$ puts positive probability on negative values of τ, and that the realizations of τ are not independent across the firms that are active in a given period.

The example used in the computations puts

$$(6) \quad \tau = v_1 - v,$$

where,

$$v_1 = \begin{cases} 1 \text{ with probablity } ax/(1+ax) \\ 0 \text{ otherwise,} \end{cases}$$

and,

$$v = \begin{cases} 1 \text{ with probability } \delta \\ 0 \text{ otherwise.} \end{cases}$$

By making the time period per decision small relative to the time period in the data we can generate distributions of increments that make large changes in τ possible (in which case we would want to adjust the other parameters of the problem, eg. the discount rate, accordingly).

To choose optimal investment and exit policies incumbents need also a perceived distribution for the number and efficiencies of their future competitors, for

$$(7) \quad \hat{s}_{t+1} = s_{t+1} - e[\omega_{t+1}],$$

where $e[\omega_{t+1}]$ is a vector which puts one in the ω_{t+1} spot and zero elsewhere, conditional

on s_t, and ω. This will be denoted by

$$q_\omega\{\hat{s}_{t+1}|s_t\} = \Sigma q_\omega\{\hat{s}_{t+1}|s_t,v_{t+1}\}p\{v_{t+1}\}.$$

Note that this distribution embodies the incumbent's beliefs about entry and exit.

We assume that the functions $q_\omega[\hat{s}|s]$ can be derived as the transition probability for $s_{t+1} - e[\omega_{t+1}]$ from some regular Markov transition kernel, say $Q[\cdot|\cdot]:SXS\rightarrow[0,1]$ and that S is compact. Ericson and Pakes show that the Markov–Perfect Nash equilibrium will generate transition kernels with these properties (i.e. that these conditions are indeed satisfied in equilibrium).

Given that $q_\omega(\cdot|s)$ provides the incumbents perceived distribution of future market structures, the Bellman equation for the firm's maximization problem is

$$(8) \quad V(\omega,s) =$$
$$\max \{\Phi, \sup_{(x\geq 0)} [\pi(\omega,s) - cx + \beta \Sigma V(\omega+\tau,\hat{s}+e[\omega+\tau]) \, q_\omega[\hat{s}|s,v] \, p(\tau|x,v) \, p(v)]\}.$$

Φ is the sell off value of the firm. If it is greater than the expected discounted value of net cash flows from staying in operation conditional on optimal behavior in each future period (the second argument after the max operator), the firm shuts down.

Ericson and Pakes (1989) provide a reasonably detailed exposition of the nature of optimal policies in this framework. What we require here is the fact that boundedness of the value function implies that if ω is high enough, the value of additional increments in ω can be made as small as we like. Since the return to investment in this model is determined by the increment in the value function generated by higher values of ω, the boundedness assumption insures that investment will be zero for all ω greater than some $\hat{\omega}$. Since firms cannot improve their quality index without some investment, states above $\hat{\omega}$ are "coasting states" from which the firm's ω can only deteriorate (and will stochastically).

So there is an upper bound to the achievable ω states (this argument is conditional on $\hat{s}$, but since S is compact, there is an overall upper bound). Similarly, the possibility of exit generates a lower bound for the observed ω states. So we can, without loss of generality, assume $\omega \in \Omega = \{1,...,K\}$.

To complete the description of the model we need also to specify the primitives which determine entry behavior. We have chosen a very simple model of entry where entry is sequential from an unlimited pool. Here a potential entrant who decides to enter pays a (sunk) setup fee of $x^e(m)$, which is nondecreasing in the number of entrants (m), then obtains a draw from the distribution $P[\omega_0]$ and begins operation in the next period at the ω–location generated by that draw. Potential entrants enter if the expected discounted value of net cash flow from entry exceeds $x^e(m)$.

Formally, if $q_{m-1}[\hat{s} \mid s,\nu]$ provides the perceptions of the mth potential entrant of the distribution of future market structures and

$$V^e[s,m] = \beta \, \Sigma \, V[\omega,\hat{s}+e_\omega]q_{m-1}[\hat{s} \mid s,\nu]p[\omega_0 \mid \nu]p(\nu)$$

then

$$(9) \quad m_s = \begin{cases} 0 \quad \text{if} \quad V^e[s,1] \leq x^e_1 , \quad \text{else} \\[2mm] \min \{m \in \mathscr{Z}_+ : \quad x^e_m \leq V^e_{s,m} , \, x^e_{m+1} > V^e_{s,m+1}\}. \end{cases}$$

Note that the distribution of entering ω's is fixed over time. Thus the "ability" of entrants progresses at the same pace as the "ability" of the outside world (in terms of our example it advances with the ability of the outside alternative). If this were not the case entry would eventually go to zero and stay there. Also in the computational example we set $x^e_1 = x^e$, and $x^e_2 = \infty$, so the maximum number of entrants in any given period is one (the maximum number of entrants in any time interval depends on the number of decision–making periods in that time interval), and

$$(10) \quad p[\omega_0 | v] = \begin{cases} 1 \text{ for } \omega_0 = \omega^*, & \text{if } v = 0 \\ 1 \text{ for } \omega_0 = \omega^* \text{-}1 & \text{if } v = 1. \end{cases}$$

Ericson and Pakes first show that under our conditions there is an upper bound to entry in any period. That is for all s , $m_s \leq m^*$. They then show that there exists an industry size (say N^*) such that whenever there are N^* or more incumbents there will be no more entry. That is if $\Sigma s_j \geq N^*$, then $m_s = 0$. Hence there will always be less than $N^* + m^*$ firms active in a given period, and since each must have an $\omega \in \{1,...,K\}$, the cardinality of S (the set of attainable industry structures) is less than $K^{N^* + m^*}$; i.e. there is only a finite number of industry structures possible.

They also provide a formal definition and proof of the existence of a Markov Perfect Nash equilibrium under these assumptions. In equilibrium, firm behavior (incumbent and entrant) depends on the perceived distributions of industry structures formalized in the transition functions, $q(\omega)[\cdot | \cdot]$. Yet the investment, entry, and exit choices, generated by that behavior, together with the known distributions of τ given alternative values of x, and $P[\omega(0)]$, generate an objective distribution of industry structures. The model is considered consistent for a $Q(s' | s)$ if and only if the objective distribution generated by the investment, entry, and exit decisions which result from a perceived distribution of industry structures equal to $Q(s' | s)$, is $Q(s' | s)$. Any such $Q(s' | s)$ (and there may be more than one of them) will be associated with a rational expectations, Markov Perfect, Nash equilibrium.

The industry structures generated by this equilibrium will all be counting measures on $\mathfrak{N}$ with a finite number of firms, i.e.

$$S = \{ s = [s_1,...,s_K] : \Sigma s_j < N^* + m^* \},$$

with the vector $[s_1,...,s_K]$ providing the number of firms at each possible efficiency level. So the heart of the equilibrium is a stochastic process for industry structures [for $\{s_t\}$], defined on $(S^\infty, \underline{S}, \underline{P})$. This process is Markov i.e., if $s^t = (s_t, s_{t-1},...,s_1)$, then

$$\Pr\{s_{t+1} = s' \mid s^t\} = \Pr[s_{t+1} = s' \mid s_t] \equiv Q[s' \mid s_t],$$

with transition kernel $Q[\cdot \mid \cdot]$, and initial condition s_0 (assumed in S). The Ericson–Pakes paper also proves that this transition kernel is ergodic. We now provide first a technical, and then a verbal, description of the ergodic result.

Formally it is shown that: the state space, S, contains a unique, positive recurrent, communicating class, say $R \subseteq S$; that there exists a unique probability measure, say μ^*, whose support is R, and which satisfies, $\mu^* Q = \mu^*$; and that if $\mu_T[s_0, s]$, gives the fraction of time periods for which $s_t = s$, then $\mu_T[s_0, s] \rightarrow \mu_s^*$ a.s. uniformly over $s \in S$.

In words we have that no matter s_0, s_t will eventually wander into a subset, say R, of the possible industry structures. Once in this set there is no probability of communicating with states outside R (all other states are transient). The states in R all communicate with one another (with finite expected transition times), so every point in R is visited infinitely often. Eventually, the sample frequencies of landing on the various points in R goes to μ^* (a vector which is reproduced by the transition kernel $Q[\cdot]$). Consequently sample averages of all random variables which are functions of s will converge to their expected values with respect to the probability distribution given by μ^*.

Note that though $(1/T)\Sigma s_t \rightarrow \mu^*$, s_t itself never settles down. Rather the structure of the industry is in perpetual flux. Depending on the nature of $Q[\cdot]$ we can expect the industry to go through periods when output is concentrated in the hands of a small number of large firms, and then, perhaps as a reaction to a sequence of new inventions, to fracture into an industry composed of a large number of approximately equally sized firms. Of course even over periods when the industry structure remains

relatively stable there will be heterogeneity in the outcomes of the active firms, with rank reversals, and simultaneous entry and exit as the normal course of affairs.

It is worth emphasizing, however, that the actual nature of the limit distribution, i.e. of μ^* (whether in fact it does include both relatively fractured and relatively concentrated structures), and the nature of the pattern of likely transitions between elements in that limit distribution (do we cycle over the divergent types of structures, or are there Poisson type events that take us more directly from one type of structure to another), depends on the nature of the primitives of the model; the profit function, the discount rate, the sunk costs of entry and exit, and the impact of investments. These in turn, depend on demand patterns, technological opportunities, and the institutional structure of the industry.

What the ergodic theorem tells us is that if we are willing to suffice with limit properties we can analyze them, and how they react to different values of the parameter vector, without specifying initial conditions. It is silent on both the characteristics of the limiting averages, and on the likelihood of the alternative sample paths that could lead to them. On the other hand, if we could compute the transition kernel, $Q[\cdot \mid \cdot]$, we could analyze the impact of alternative events from the actual initial condition prevalent in the economy, and compute the ergodic distribution of industry structures, together with both their welfare and descriptive implications, directly.

II. Computational Algorithm

The model has been set up so that most of its parameters can be estimated without actually solving for the fixed point which defines the stochastic dynamic equilibrium. Indeed one of our intentions in setting out an internally consistent dynamic model was to provide a framework for evaluating the properties of alternative estimators when the agents that generate the data are forward looking and do interact with one another (for an

example of this type of application, see Olley and Pakes,1991).[6] Though that may well turn out to be the major advantage of our approach, we also wanted to be able to provide a complete analysis of the equilibrium process and its relationship to the primitives of the model. In particular, we wanted to be able to trace out how the industry's equilibrium path was likely to respond to changes in policy or in the environment. To do this we need a computational algorithm that allows us to solve for (or simulate) the stochastic process generating $\{s(t)\}$ for different values of the parameters of the model.

We turn now to a description of one such computational algorithm, and its limitations. We then use that algorithm, together with the functional forms given in the example provided above, to compare the set of industry structures that would be generated from one particular set of technological parameters under alternative institutional regimes. The reader who is not interested in the computational detail but might be interested in the results of the numerical analysis should be able to omit this section and have no difficulty with section III.

What we need is the form of $Q[\cdot \,|\, \cdot]$. This transition kernel is computed from the transition probabilities for all incumbents and potential entrants, and these in turn are functions of the optimal policies and the primitives of the model. Though the actual computation may well be complicated (see Ericson and Pakes,1989), given the optimal policies (entry and exit rules and the investment policies of incumbents), and an initial s, it is easy to simulate different sample paths and then derive an empirical distribution which will converge (uniformly) to the true $Q[\cdot \,|\, \cdot]$. So the whole computational problem is in

[6]The estimation problem breaks down into subproblems which, provided care is taken to account for the simultaneity of input decisions and the selection induced by entry and exit, allow one to obtain estimators of subvectors of the model's parameters. The natural breakdown is to esimate the parameters of the cost and demand functions that determine the spot market equilibrium in one part of the analysis, and to use data on the evolution of market shares and investment expenditures to estimate the parameters of the family $\mathbb{P}$, and β. This still leaves us without estimators for the sunk costs of entry and exit, which will either have to be gotten exogenously or estimated from the implications of the more complete framework.

finding the optimal policies. To obtain these policies we compute the value of being active for different values of a firm's own and of its competitors state variables. This is a map from $\Omega \ X S \to \mathbb{R}$. We turn now to a description of a computational algorithm designed to find this function and the associated polices and a discussion of its properties.

A. Computational Algorithm.

We work off value functions for problems with a limited amount of active firms, and then push that limit up. Start with the value function for 1 active firm. This is a straightforward contraction which sets the support of (the upper and lower limits for) ω, i.e. it sets Ω.

For N=2 we need to calculate $V[\omega_1;\omega_2]$ for $\{\omega_1,\omega_2\} \in \Omega^2$. Start with $V^0[\ ,\ ] = \pi[\omega_1;\omega_2]$. Then get $V^k[\omega_1;\omega_2]$ as the solution to:

$$(1a) \quad V^k[\omega_1;\omega_2] = \pi[\omega_1;\omega_2] - cx + \beta \Sigma_{\tau_i,\tau_j,v} V^{k-1}[\omega_1+\tau_1+v;\omega_2+\tau_2+v]p[\tau_1|x_1,v]p[\tau_2|x_2,v]p(v),$$

where the couple $[x_1,x_2]$ satisfy the Kuhn–Tucker conditions

$$(1b) \quad x_i\{-c + \beta \Sigma_{\tau_i,\tau_j,v} V^{k-1}[\cdot;\cdot]p_j(\partial p_i/\partial x_i)p(v)\} = 0,$$

$$x_i \geq 0,$$

for $i \neq j$, $p_i = p[\tau_i|x_i,v]$, and so on. One way of calculating the fixed point would be to iterate on this procedure until $\|V^k-V^{k-1}\|$ was below an acceptable tolerance for an appropriate norm $\|\cdot\|$.

This procedure does away with one of the "nests" that is implicit in the fixed point calculation. To see this note that a straightforward alternative way of calculating $V(\cdot)$ would be to begin with a candidate function for the process generating the competitors

states [for $p(\tau_2|s,\nu)$ in our two firm example], use it to solve the implied contraction for $V[\cdot\,|p(\tau_2|\cdot)]$ and the associated investment policy, then use that investment policy to calculate a new process generating the competitors states, and iterate on this procedure. By solving for all the x's simultaneously at each iteration our procedure has done away with the second fixed point calculation at each iteration. However, the procedure in (1) does this at the cost of solving an implicit nonlinear system of equations at each iteration of the fixed point calculation, and the nonlinear search increases computational time dramatically.

An alternative is to use x_j^{k-1}, last iteration's x–value for one's competitors, in the Kuhn–Tucker condition for x_i. If we ignore the constraint that $x \geq 0$, this gives an explicit solution for each x_i. If $x_i \leq 0$, set it equal to zero. This does away with the need for solving a nonlinear system in N variables at each iteration, and replaces it with the solution to N univariate problems (each of which is nearly linear in our example). Computational time decreases accordingly. Of course we could iterate on this "policy step" until convergence and (provided convergence is achieved) solve the nonlinear system exactly without ever having to compute an explicit solution to the nonlinear system. Though it did not prove helpful to iterate on this "policy step" in the examples presented here, it is more likely to prove helpful once we substitute a polynomial approximation into the computational algorithm(see below).

For expositional simplicity we ignored entry and exit in the discussion above. To account for exit we simply compare $V^k(w_1;w_2)$ to Φ. If the latter is larger we set $V^k()=\Phi$ and the associated x to zero. To account for entry, consider states where $V[w_2,w_1] = \Phi$. In such cases we use $\pi[w_1]$ instead of $\pi[w_1;w_2]$, and calculate $V^e[w_1]$. If this term is greater than x^e, we calculate $V[w_1;w_0]$ for $V[w_1;w_2]$, where the transition probabilities for w_0 are given by $P[w_0]$.

We still continue iterating until $\|V^k-V^{k-1}\|$ is within a given tolerance, but now also check $\|x^k-x^{k-1}\|$. A fixed point to this problem can be shown to satisfy all the

requirements for our equilibrium if the maximum number of firms (our N), set either endogenously by the model's parameters, or by an artificial barrier to entry, is 2. Now push N up to 3 and do the iterative calculation again starting at

$$(2) \quad V^{k=0,N=3}{}_{[\omega_1;\omega_2,\omega_3]} = V^{k=\infty,N=2} \{\omega_1;\max[\omega_2,\omega_3]\}.$$

This procedure should be repeated until we reach an N where, for all industry structures at which $\Sigma s(i)=N$, $V^e(s,1)<x^e$. The N which satisfies this condition is an upper bound to the number of firms ever in the industry (provided, of course, that the initial s has no more than N firms).

Some caveats are in order before proceeding. We have not proved a contraction property for our algorithm yet, so we have no way of knowing for sure whether it converges. On the other hand it has converged for every set of parameter vectors we have tried. We have not proven uniqueness either. We did calculate all results we describe here thrice, starting each at different initial conditions [once at $\pi(\cdot)$, once at zero, and once at $V(\cdot)$ from the smaller N, see above]. In each case we got to precisely same answer.

III. Numerical Illustrations.

To illustrate the type of equilibria that result from this class of dynamic heterogeneous agent models, we have computed the policy functions, and simulated the equilibria, for the following set of parameter values:

δ (the probability that the outside alternative moves up) = .7;

β (the discount rate) = .925,

x^e (sunk entry cost) = .2,

Φ (scrap value) = .1,

M (size of market) = 5,

a = 3 (determines the efficacy of own investment in

increasing the probability of quality improvements),

mc (marginal cost of production) = 5,

$$e^{g(\omega)} = \begin{cases} e^{\omega} & \text{if } \omega \leq \omega^*, \\ e^{\omega^*}[2 - e^{-(\omega - \omega^*)}] & \text{otherwise.} \end{cases}$$

The total current cost of producing the output of this industry is about 25, so the sunk entry cost, i.e. x^e, is about $1/125^{th}$ of the cost of production in a given period. That is, it is not very costly to enter in this, our base, case. An arbitrary unit can be attached to M (i.e. thousands, millions,...), and the marginal cost is in those units (the cost of producing a thousand or a million units,...). The functional form for $g(\omega)$, in particular the difference between $g(\omega)$ and ω for sufficiently large ω, is to insure that profits are a smooth and bounded function of ω.

The columns labelled "MP" in the tables to follow were constructed as follows. We substituted these parameter values into the computational algorithm given in the last section, and let that algorithm compute the Markov Perfect Nash investment, entry, and exit policies. We then used these policies, together with a pseudo random number generator, to generate ten thousand periods of industry evolution from an initial condition of an industry with one firm at $\omega = \omega_0$. The other columns in these tables will be described in more detail presently. They were constructed in a similar way to the MP column but with a change in either a behavioral assumption, or in a parameter value. In particular for the column labelled "$\sigma^* = .65$" we imposed an institutionally created upper bound to the market share of the largest firm; in the column labelled "$x^e=2$" we imposed an institutionally created tenfold increase in the sunk cost of entry; in the column labelled "Coll" we used the parameters above but assumed there was a perfect cartel controlling all pricing, investment, entry and exit decisions; and in the column labelled "PP" we used the

parameters above but assumed there was a benevolent social planner controlling all pricing, investment, entry and exit decisions.

Table 1 provides some statistics which help describe the evolution of market structures in the ergodic distribution of market structures that the various solutions generated. We will focus on the Markov Perfect column first. Part A of the table indicates that the ergodic process characteristically has either three or four firms active in a given period. There is, however, lots of entry and exit, so the firms active in equilibrium are not always the same three or four firms. Note also that entry and exit are positively correlated; in most years when there is entry there is also exit (Part B of the table). This is in stark contrast to models of industry dynamics that do not allow for idiosyncratic sources of change, for these models characteristically predict either entry, or exit, but not both, in the same period.

The firm specific, or idiosyncratic, sources of uncertainty also generate a significant amount of job turnover (even though total demand for the industry's products barely changes from period to period). We calculated gross job creation and gross job destruction figures by assuming that the new jobs available during a year were proportional to the increases in output in firms that increased their output over that year, and the jobs destroyed during a year were proportional to the fall in output in firms that decreased their output over the year. On average about 4% of the jobs were destroyed each year (with approximately the same percentage of new jobs created during the year); but that figure was surprisingly variant over subperiods (its standard deviation was about twice its mean). So there were periods in which the industry was undergoing a lot of flux, or churning, and periods in which it was more stable (see below).

About one thousand and eight hundred firms participated in the industry during the 10,000 periods simulated (Table 2, columns MP). Most, however, remained active only a short period of time. The modal lifespan was only one year (period), and the median was about the same. On the other hand the distribution of lifespans was extremely skewed

Table 1					
Characteristics of Ergodic Distribution[1] $\delta = .7 \quad \alpha = 3 \quad \beta = .925 \quad xe = .2 \quad \$ = .1$ $m = 5 \quad c = 5$					
No. of Time Periods	10,000				
	MP	$\sigma^* = .65$	$x^e = 2$	PP	Coll.
% with 6 firms active	.1	.2	.0	0	0
% with 5 firms active	1.6	3.1	.3	.1	.1
% with 4 firms active	35.3	33.3	1.2	5.8	1.1
% with 3 firms active	63.0	63.3	17.1	44.5	23.1
% with 2 firms active	.0	0.0	81.4	49.6	75.6
% with entry and exit	13.1	11.5	.3	10.1	10.7
% with entry only	4.8	4.8	.7	3.0	1.9
% with exit only	2.0	2.5	.6	2.3	1.7
% with entry or exit	20.1	18.7	1.6	15.4	14.2
Gross job creation	.086	.086	.031	.033	.027
Gross job destruction	.087	.088	.032	.033	.027

1. Legend

    MP   = Markov Perfect Nash Equilibrium
    σ^* = .65, MP with market share constrained to be below .65.
    x^e = 2, MP with sunk entry costs increased to 2.
    PP   = planner's problem (see the text below).
    Coll. = perfect cartel (see the text below).

Table 2[1]

Lifetime Distribution

(Based on 10,000 time periods)

$\delta = .7 \quad \alpha = 3 \quad \beta = .925 \quad xe = .2 \quad \Phi = .1$
$m = 5 \quad c = 5$

Life time	Percent			Implied Hazard			Cumulative Percent		
	MP	$x^e{=}2$	PP	MP	$x^e{=}2$	PP	MP	$x^e{=}2$	PP
1	48.3	17.5	25	48.3	17.5	12.2	48.3	17.5	25.0
2	24.7	14.6	44.0	47.7	19.6	24.0	73.0	35.0	68.9
3	6.1	4.7	9.0	22.6	7.3	17.9	79.1	40.7	77.9
4	3.6	8.5	4.4	17.2	14.3	17.9	82.7	49.1	82.3
5	2.5	3.8	1.9	14.4	7.5	10.9	85.2	52.1	84.2
6	1.3	2.8	2.2	8.78	5.9	8.9	86.5	55.7	86.5
7	.8	1.9	1.4	5.9	4.5	4.1	87.3	57.2	87.9
8	.7	.9	.8	5.5	2	4.8	88.1	58.5	88.6
9	.7	.9	.5	5.9	2.9	5.7	88.8	59.5	89.2
10	.3	1.9	.7	2.7	4.6	5.7	89.1	61.3	89.9
	Number Active			Mean			Median		Standard Deviation
MP	1,800			18.4			2		73.2
S	103			191.5			6		428.0
PP	1301			19.5			2		93.9

1. See the note to Table 1.

with a mean of 18.4 years, and a standard deviation of 77.5 years. Those firms which did survive the initial period went on to be part of the industry for long periods of time. Relatedly, both mortality and hazard rates decline markedly over the first seven or eight years, giving the impression that the initial period is a "learning" period (recall that the hazard rates shown in the table are estimates of the true hazard rates implied by the model; their standard errors being on the order of .005).

Table 3 provides characteristics of the realized values of the firms which participated. The first point to note is that over 90% of the firms which participated in this industry had a net loss from their endeavor (generated negative realized values). Most lose about .2 (the entry fee), but there are those who invest for awhile, never move up the "quality ladder", and eventually drop out, losing also their investments in the interim. Among the 9.4% whose realized values were positive, the mean realized value was a very high 10.1 (recall that the sunk cost of entry was only .2 so this implies a benefit/cost ratio of 50.5), and the distribution of realized values was very skewed (about 3% of them earned 57% of the total distribution of positive realized values).

These parameter values generate an industry in which it is relatively cheap to start up and explore some new idea. Most startups are not successful. The few that are successful earn, on average, very high rates of return, and even among them the distribution is very skewed, so that the 2 or 3% most successful earn rates of return that are really quite phenomenal. Of course eventually even the most profitable are taken over by better ideas and find it optimal to exit.

This pattern of high initial mortality rates, and skewed returns and lifespan distribution for those few firms who survive the initial high mortality period, has an analytic counterpart in the shape of the value function for this problem (that function can be computed directly from the output of the algorithm described in the last section). If we hold the w's of all competitors constant and compute a section of the value function showing how it changes with changes in the firm's own w, we find it to be initially convex

Table 3

(see legend on Table 1)

Realized Value Distribution

$\delta = .7 \quad \alpha = 3 \quad \beta = .925 \quad xe = .2 \quad \Phi = .1$
$m = 5 \quad c = 5$

Obs/Num	Realized Values			Life Time	Sum of Realized Values		
	MP	$\sigma^* = .65$	$x^e = 2$	MP	MP	$\sigma^* = .65$	$x^e = 2$
1	72.8	45.1	83	79	72.8	45	83.0
2	52.6	38.9	34.0	247	125.4	84	117.0
3	33.1	36.3	32.4	718	158.57	120.3	149.4
4	32.7	35.7	28.4	118	191.3	156	177.8
5	29.3	30.2	20.2	102	220.6	186.3	198.0
10	22.8	21.2	15.74	5	343.8	301.3	282.6
100	7.11	6.8	−4.43	215	1462.6	1313.9	294.6
150	1.81	1.1	\	37	1700.8	1508.5	\
170	.1	−.06	\	3	1717.6	1514	\
171	−.05	−.07	\	2	1717.5	1514	\
1491	−.10	−.7	\	4	1586.5	1336.1	\
1800	−4.08	−	−	15	1282.05	−	−
	Mean	Median	Std.Dev	# Positive	Mean of Positive	#Negative	Mean of Negative
MP	.71	−.1	4.09	170	10	1630	−.27
$\sigma^* = .65$	.68	−.1	3.74	150	9.5	1465	−.28
$x^e = 2$	2.37	−1.89	11.28	329	14.2	74	−2.5

and then concave in ω (for a proof of this property for a simpler problem, see Ericson and Pakes,1989). Since the marginal benefits of investment is increasing in the slope of this value function, we expect little investment at the initial low values of ω. Recall that the change in ω is stochastically increasing in investment, so the low investment levels will generate a downward drift in ω at low ω-values, and hence high mortality rates. The odd entrant whose ω does take a sequence of steps upward will begin to invest more, and continue to increase its investment, with the consequent improvement in the distribution of the increment's in the firms's ω, until its ω passes the inflection point. From then on increases in ω will be accompanied by falling investment. However, should a firm with high ω drop down to near the inflection point, its investment will start increasing again. As a result mortality probabilities for a firm which has developed its ω past its inflection point are very low, and a firm which has become successful is likely to be a profitable actor in the industry for some time. Note that any problem in which a firm has to develop its product somewhat before that product can be successfully marketed, and in which profits are bounded, will generate an initial convex, and a final concave, portion to the value function. So there are a broad variety of functional forms for which we ought to expect high initial mortality rates, reasonably long lifespans conditional on survival in the high mortality period, and "coasting" states for very successful firms.

Recall that there is free entry into this industry. As a result, apart from integer problems, entrants earn, on average, discounted returns that are just equal to the sunk cost of entry (actually about .26). Of course if one limits themselves to either rates of profit, or market values, of *active* firms one obtains an indication of supernormal rates of return (an average per period profits of 5.42, and an average market value of over 20). This is, of course, a simple result of the fact that the observed data is dominated by firms which survived for long periods, and *they* did indeed earn very high rates of return on their sunk investments. That is we ought to expect supernormal average returns for active firms even in the most "competitive" of industries; for the observed data is selected on the basis of

success.

Tables 4A and 4B (again column MP) provide summary statistics on the distribution of the one firm concentration ratio, and on price cost ratios, over the 10,000 periods, while figure 1 plots the evolution of the one firm concentration ratio over a period. For ninety percent of the periods, the average percentage markup lies between 27 and 44 percent. Accordingly, the industry is most often reasonably fractured. The one firm concentration ratio averaged .39 in an industry in which there is almost always either three or four active firms. Periodically, however, a firm will surge ahead of its competitors and stay there for long periods of time, creating epochs where the industry looks quite concentrated. For example there is an interval of forty periods between 320 and 360 when the concentration measure never falls below .4, averages about .6, and even goes through a 10 year stint at .8. For the next 50 periods the concentration ratio never goes above .4 and is almost always between .3 and .35. Note that these changes in industry structure occur quite naturally, without any change in the external environment.

A. "Policy Experiments".

Once we have estimates of the relevant parameters, and the ability to compute the equilibria that result from them, it is easy to perform a host of experiments. This subsection provides some illustrations. We stress, however, that they are just that —illustrations. There is little empirical basis for picking the parameters we do, and, not much analytic knowledge of whether the qualitative results persist under alternatives. What we are providing is a series of examples of what can happen.

We consider a set of perturbations to the equilibrium just described designed to investigate the impacts of institutional change on both market structure, and on welfare. The welfare associated with each different problem is calculated as the discounted sum of the cash flows accruing to the active firms (profits minus investment), plus the discounted sum of exit fees minus entry costs, plus the consumer surplus generated by the products in

		MP	$\sigma^* = .65$	$x^e = 2$	PP	Coll
.95	quantile	.62	.61	.63	1.0	1.0
.90	quantile	.5	.5	.63	.84	1.0
.75	quantile	.43	.43	.51	.52	.72
.50	quantile	.34	.34	.5	.5	.5
.25	quantile	.33	.33	.5	.46	.5
.10	quantile	.26	.3	.34	.33	.33
.05	quantile	.25	.25	.46	.33	.46
mean		.39	.38	.51	.54	.60
standard deviation		.11	.097	.09	.20	.20

Table 4A[1]

One - Firm Concentration Ratios

1. See the note to Table 1.

	MP	$\sigma^* = .55$	$x^e = 2$	Coll.
		Table 4B[2]		
		Price / Cost Ratios		
MAX	2.11	2.21	2.18	2.45
.95	1.44	1.48	1.46	2.34
.90	1.40	1.39	1.46	2.33
.75	1.33	1.33	1.40	2.33
.50	1.30	1.30	1.40	2.33
.25	1.30	1.30	1.40	2.24
.10	1.27	1.27	1.30	2.22
.05	1.27	1.27	.3	2.20
Min	1.24	1.23	1.27	1.29
Mean	1.33	1.32	1.41	2.30
Standard Deviation	.085	.073	.099	(.089)

1.   See the note to Table 1.

Share of Largest Firm
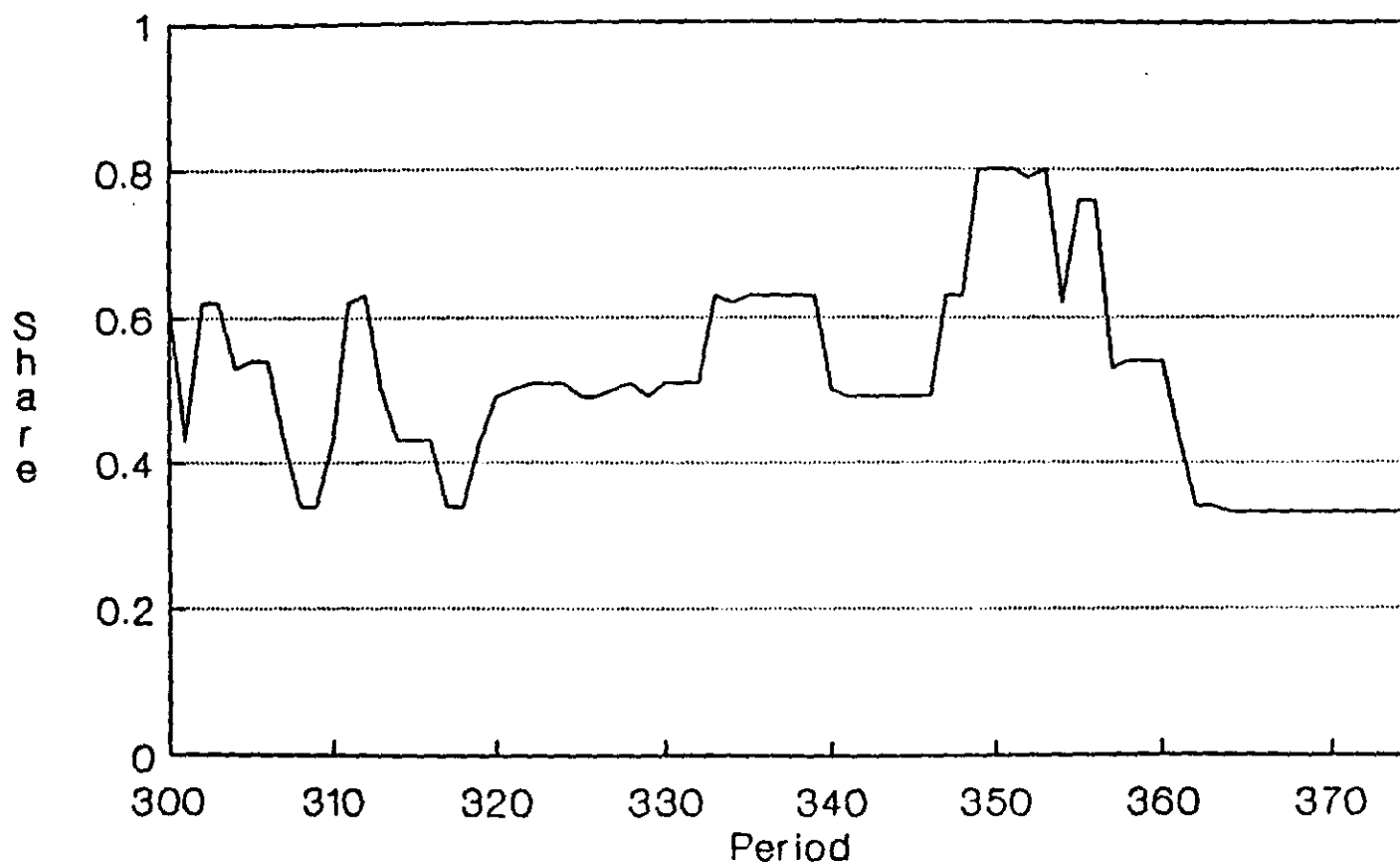
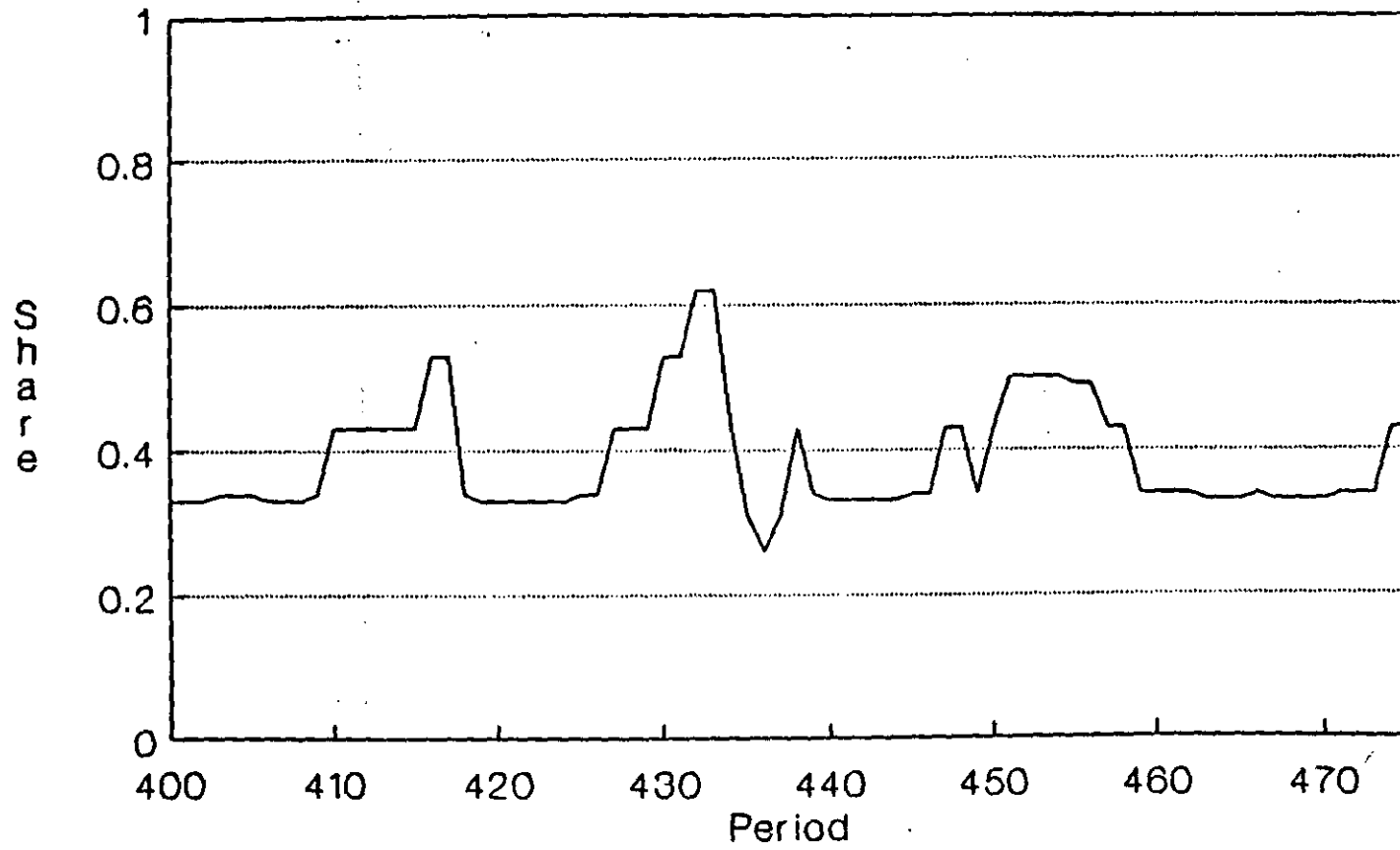$\delta = .7 \quad a = 3 \quad \beta = .925 \quad xe = .2 \quad \phi = .1$

$m = 5 \quad c = 5$

Share of Largest Firm

$$\delta = .7 \quad a = 3 \quad \beta = .925 \quad xe = .2 \quad \delta = .1$$

$$m = 5 \quad c = 5$$

existence in each period. The latter is calculated as

$$(1) \quad \int [\max(\omega_j - p_j + \epsilon_j)]dG(\epsilon_1,...,\epsilon_n) = \log \Sigma \exp(\omega_j - p_j)$$

where the equality is a result of the fact that $G(\cdot)$ is multivariate extreme value (see McFadden, 1984), and n is the number of active firms in the period.

Recall that for any given institutional environment there will be a distribution of welfare results (the outcome will depend on the realizations of the investment processes). As a result we simulate 100 runs for each institutional environment, and then list the resulting means and standard deviations for; net cash flow to the firms, consumer surplus, and total welfare.

The initial condition for each institutional environment is an industry started by a single firm with a new product at an $\omega = \omega_0$; i.e. $s_0 = e(\omega_0)$. In the first example we bestow monopoly power on the initial entrant for T periods, and allow the free entry Markov Perfect equilibrium with initial condition $s_{T+1} = e(\omega_{T+1})$ thereafter. Next we consider a variant to this case, in which the monopolist is allowed to produce as many products as it likes. In this case, which we call the perfect cartel, a single decision maker controls the pricing, investment, entry, and exit decision of each firm (or product). In the third case we assume an "entry barrier" in the form of a licensing fee which increases the sunk cost of entry by a factor of ten, from .2 to 2. The fourth case can be viewed either as a regulatory constraint, or as a reaction of firms to the possibility of future regulatory constraints. In this case we assume that the firm is constrained to keep its market share below some σ^*. Each of these alternatives is compared both to the original Markov Perfect equilibrium, and to the market structures and welfare that would be generated by a benevolent social planner.

A different (and much simpler) algorithm is used to calculate the policy functions for the cartel and the social planner. Let $B(\omega_1,...,\omega_n)$ be the one period (or current)

benefits to the decision maker when there are n firms active and their efficiencies are $(\omega_1,...,\omega_n)$. The social planner is interested in maximizing social surplus, so it sets each price equal to the common marginal cost, and then calculates the one period benefits from the consumer surplus calculation in (1) above. The cartel is interested in maximizing the discounted sum of the total profits of all firms in the industry. In each period it chooses a vector of prices to maximize the total profits that could be generated from the n goods currently marketed and calculates

$$(2) \quad \max_{(p_1,...,p_n)} \Sigma_j \, M \, \sigma(\omega_j,p_j;\omega_1,...,\omega_n,p_1,...,p_n) \, [p_j - mc],$$

where $\sigma(\omega_j,p_j;\omega_1,...,\omega_n,p_1,...,p_n)$, the market share of the jth good given the prices and efficiencies of all goods marketed, is calculated as in (I.3) above.

The value function for the planner and the cartel are calculated from the following recursion. If $V^{k-1}(\omega_1,...,\omega_n)$ is the k–1 iteration of this function, and we have ordered the ω so that $\omega_i \geq \omega_{i-1}$, then we calculate the kth iteration value function as

$$(3) \quad V^k(\omega_1,...,\omega_n) = \max_{(1 \leq i \leq n)} V^k(\omega_1,...,\omega_i)$$

where

$$V^k(\omega_1,...,\omega_i) = B(\omega_1, \quad ,\omega_n) + (n-i)\vartheta + \max$$

$$\{\sup_{(x_1,...x_i)} - c\Sigma_j x_j + \Sigma_j V^{k-1}(\omega_1',...,\omega_i')p(\omega_1'|\omega_1,x_1,\nu)...p(\omega_i'|\omega_i,x_i,\nu)p(\nu);$$

$$\sup_{(x_1,...x_i)} - c\Sigma_j x_j - x^e + \Sigma_j V^{k-1}(\omega_1',...,\omega_i',\omega^e)p(\omega_1'|\omega_1,x_1,\nu)...p(\omega_i'|\omega_i,x_i,\nu)p(\omega^e|\nu)p(\nu) \}.$$

If the first max over i for $i\epsilon\{1,...,n\}$ is i*, then (n–i*) of the incumbents exit. The second max operator determines whether or not a new product is introduced in the following period.

Table 5 provides the welfare results. We begin by comparing the various

Benefits/ Market Structure	Total Firm Cash Flows		Consumer Benefits		Total Benefits[**]	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Monopoly	207	66	96	17	303	83
20 Year Monopoly then free entry	180	57	140	21	320	74
10 Year Monopoly then free entry	146	54	186	31	331	71
Markov-Perfect Nash	70	26	301	65	369	68
Perfect Collusion	218	55	115	19	332	74
Social Planner	/	/	/	/	377	/
$\sigma^* = .65$	61.7	15.1	289.7	64.4	349.6	67.4
$\sigma^* = .55$	54.4	12.0	284.8	66.1	337.5	73.1
Sunk Costs = 2	76.5	26.3	293.4	55	361.5[***]	69.8

Table 5
Social Welfare From
Alternative Market Structures[*]

*All calculations are based on 100 simulation runs and the parameter values in Table 1. Each simulation run uses the initial condition that there is only one firm active and it is at ω_0.

**Total Benefits are net cash flows plus consumer benefits minus the discounted value of entry fees minus exit fees (the latter only make a small contribution to the calculation.)

***The mean of the discounted incremental entry fee is 7.8, when $x^\ell = 2$.

temporary monopolies to the Markov Perfect Nash equilibrium. The columns with the means of the discounted cash flows, and of the discounted consumer benefits, show that, as expected, there is a large distributive effect of going from one institutional environment to another. An infinitely lived monopolist will earn on average about three times as much as the total discounted net cash flows of all the competitors in a Markov Perfect Nash equilibrium. However, consumer surplus will be about three times as high when we allow for free entry. Since the two impacts of monopoly on welfare tend to offset one another, the total impact on welfare is not as dramatic. Still monopolization does decrease welfare by about 20%.

At least two caveats should, however, be stressed here. First these institutional arrangements restrict the monopolist to produce no more than one product. Though, depending on the source of monopoly power, there may be some limits on the number of goods a monopolist can market, a limit of one is somewhat artificial. The extreme alternative is to assume that the monopolist can market as many goods as it likes — but no one else can enter the industry. The monopolist would then become a perfectly colluding cartel; a cartel that controlled pricing, investment, and entry and exit decisions to maximize the sum of the expected discounted values of future net cash flows of all of its members.

The results of the welfare calculations for this institutional environment are given in the row labelled "perfect collusion" in Table 5. As compared to the single product monopolist, producer surplus, of course, goes up; but the increase is under 5%. Consumers, however, now have a choice of products, and this increase in product diversity implies that the change in institutional structure also generates an increase in consumer surplus; in our case an increase of almost 20%. So the cartel, or the monopolist who is allowed to market many products, generates a measure of total welfare that is only 10% lower than the total welfare generated by the Markov Perfect Nash solution.

The other point to note here is that the standard deviation of the welfare gains in

any of the institutional regimes is about 20% of the average welfare gain in those regimes. This is about twice as large as the average difference in welfare between the collusive and the Markov Perfect regimes, and about the same size as the difference in welfare between the single product monopolist and the Nash solutions. On average, then, differences in welfare within any institutional structure are likely to be at least as large as differences between institutional structures. Such large standard errors provide one good reason for eschewing case study approaches to analyzing the costs and benefits of different institutional structures (and, this, even if the case studies compared were laboratory perfect in the sense that we could insure that technology was exactly the same in the different cases compared, and only the institution being studied differed).

We now compare these results to the solution to the social planner's problem. Recall that in each period the planner chooses all outputs, investments, entry, and exit to maximize social welfare given the technological constraints that are operative in the Markov Perfect Nash equilibrium.

Perhaps the most striking point to be made here is that in terms of welfare there is not much difference between the planner's solution and the Markov–Perfect Nash solution (the planner does about 2% better). That is for this set of parameters a planner could not do very much better than the Nash free entry solution. On the other hand the planner generates a *market structure* which is quite different from the Markov Perfect Nash solution; indeed in many ways the market structure generated by the planner is more similar to the market structure generated by perfect collusion (see the columns labelled PP, and Coll.in tables 1 to 4). For example, the entry and exit rates generated by a colluder are almost the same as those generated by the planner (and these are only three fourths of the rates generated by the Nash free entry solution), both the colluder and the planner generate markets where there are typically two or three firm's active whereas the Nash free entry solution typically has three or four firms active, and the one firm concentration ratio generated by the planner is much closer to that of the colluder than to that of the Nash

solution (though it does fall in between the two). Indeed the only really striking difference between the planner's solution and the collusive solution is in price cost ratios. The planner's is one, always. The colluder averages 2.3, a number markedly higher than in any of the other regimes we examine.

Table 6 sheds some further light on the relationship between the planner's, the colluder's, and the free entry Nash solution. It compares the investment and number of active firms generated by the alternative institutional environments. The planner typically markets fewer products, and generates less investment, than the Nash free entry solution. In the Nash solution the entrant and the investor do not take account of the effect of their investment and entry decisions on the profits of their competitors, while the planner does. The planner does, however, generate an equilibrium with both more firms and more investment than the equilibrium generated by perfect collusion. This because the planner takes account of the increase in consumer as well as producer surplus generated by entry (product diversity) and increases in product quality, while the colluder takes account only of the effect of entry and investment decisions on producer surplus.

Before leaving this comparison we should emphasize that our computations for the "collusive" case simply assumed that there was some outside authority that could insure that the collusive agreement was maintained (eg. a legal source of monopoly power). We have not considered whether one could support such a regime with an implicit system of punishments (recall that for this type of collusion to be self–sustaining it must also deter all but the collusively optimal entry); so there is a sense in which this is as extreme a degree of collusion as one could ever find. On the other hand it is not clear that as the degree of collusion increases the welfare measure decreases, so it might well be that less collusive institutional structures (say a regime which could control the prices and investment strategies of incumbents, but could not control entry) would generate lower welfare.

We move next to the results from the simulation experiments in which we set the

Table 6
Average Investment and Number of Active Firms
Under Alternative Institutional Arrangements

	Investment	Active Firms
Markov Perfect Nash	2.57	3.4
$\sigma^* = .65$	2.56	3.4
$x^e = 2$	1.97	2.2
Planner	1.95	2.6
Perfect Collusion	1.75	2.3

upper bound to market share, i.e. σ^*, to .65. In these runs firms choose prices to maximize their profits conditional on the price choices of their competitors *and* the market share constraint (as a result, firms whose ω would induce a market share of .65 or more if there were no upper bound, choose to increase their price and lose customers). At $\sigma^* = .65$ the market share constraint has only moderate bite; the unrestricted Markov solution has about 4% of the observations with $\sigma \geq .65$, while once we restrict market share firms at higher values of ω invest less so that, in the restricted runs, the constraint is binding just under 1% of the time. The columns labelled $\sigma^* = .65$ in the tables show that the market structures generated by the model with the market share restriction are very similar to those generated by the model without it. Indeed, it would be hard to tell the difference between the two institutional arrangements based on the observables listed in these tables. On the other hand, though the change in institutional regime has only a small effect on market structure, it does have a noticeable effect on welfare (welfare falls by over 5%). Moreover, both components of welfare fall; total firm cash flows *and* consumer surplus, so that there is not even a distributive reason for the market share restriction — at least for this set of parameters. The fall in consumer benefits probably results from the fact that the market share constraint forces firms with very high ω to raise price; as the top 5% of the price cost margins are higher when the market share restriction is imposed (Table 4B). Decreasing σ^* to .55 just moves all variables further in the same direction (now the fall in welfare is about 9%).

What happens if we go back to our base case but increase the sunk cost of entry by a factor of 10 (from .2 to 2)? Total net cash flows to producers go up somewhat, but consumer surplus goes down even more, so that total welfare falls; but only by about 3%. Indeed if we assume that the incremental sunk costs were institutionally created, had no administrative costs associated with them, and were recycled to consumers in another form, then they should be added to the total benefits in the table. When we did this calculation we found that there was no effect of the increase in the sunk cost of entry on welfare.

Though this change in sunk costs has only small effects on total welfare, it has surprisingly dramatic effects on market structure. The columns in Tables 1,2,3,and 4 that are labelled $x^e=2$ describe the market structures generated by the simulated runs when sunk costs were set equal to 2. The increase in sunk costs cuts out almost all entry (entry occurred in about 20% of the time periods in the base case but in only 1% of the cases after the sunk cost increase), and changes the correlation between entry and exit from positive to negative. It also generates an equilibrium in which new entrants last much longer (see Table 2), and the one firm concentration ratio is both much higher (it goes from .39 to .52) and much more stable over time (its standard deviation goes from .11 to .09). In sum the increase in sunk cost generates a much more stable and concentrated industry.

Note, however, that the price cost margins are not too different than what they were under our base case. Indeed the price cost margins are quite similar in all cases in which there is free entry. The relative constancy of price cost margins, together with the fact that entry and investment was higher in the free entry Nash than in the social planner's solution, provides some indication of why we would not expect welfare to fall too much as a result of the sunk cost increase. Indeed it is likely that there is a level of incremental entry costs that generates a larger value for the welfare measure than the free entry Nash solution does. On the other hand, it could not beat it by much, as the Markov Perfect Nash solution is very close to the planner's solution.

The latter two cases provide an interesting comparison. The effect of the market share restriction on welfare was large, but its effect on market structure was almost imperceptible. On the other hand the effect of the increase in sunk costs on market structure was dramatic, but its effect on welfare was almost imperceptible. It is reasonably clear then, that once we allow for the complexities of the investment, entry, and exit processes, there is no simple way to relate changes in the descriptive statistics we generally use to describe market structure to changes in welfare.

We conclude this section by reemphasizing that these results are for a set of

parameter values chosen more or less at random, and we have made no attempt to examine their robustness. The section was not written to provide a set of policy prescriptions. Rather we had three more modest goals. The first was to provide an illustration of the detail that can in fact be generated by simple dynamic equilibrium models with idiosyncratic uncertainty; detail which is a salient feature of data sets on firms. The second was to illustrate some of the likely problems with other approaches to the analysis of industry behavior. Finally, we hope we have illustrated the potential benefits from actually going to the data and obtaining reasonably reliable estimates of the primitive parameters needed to actually analyze the nature of equilibrium in any given industry.

V. Computational Burden and Approximation Techniques.

We begin with a brief discussion of the dimension of the problem that needs to be computed in order to obtain the optimal policies from the algorithm in section II, and then move on to a discussion of ways to reduce that computational burden. Our discussion will be based on the fact that the time required to compute the needed fixed point is roughly the multiple of: 1) the time per grid point evaluation; 2) the number of grid points evaluated at each iteration; and 3) the number of iterations until convergence.

An explicit calculation can be provided for the number of grid points that needs to be evaluated at each iteration. Note first that the number of distinct grid point is K^N, where N is the maximal number of firms ever active in the ergodic distribution. So the number of distinct grid points grows exponentially in the number of firms (N). Luckily symmetry, or more precisely exchangeability, implies that

$$(1) \quad V(\omega_1; \omega_2, ..., \omega_N) = V(\omega_1; \pi_2, ..., \pi_N),$$

for any N–1 dimensional vector $\tau = (\pi_2, ..., \pi_N)$ which is a permutation of $(\omega_2, ..., \omega_N)$. It follows that we need not evaluate the value function at each distinct grid point. Indeed the

number of points one needs to evaluate are the number of distinct N–element vectors with $\omega_1 \geq \omega_2, ..., \geq \omega_N$ and, $0 \leq \omega_i \leq K$, for i=1,...,N. The actual number of points that emanate from such a calculation is given by the following lemma.

2.Lemma.

The number of distinct sequences $[\omega_1, ..., \omega_N]$ with $\omega_i \geq \omega_{i-1}$, and $\omega_i \in [0,...,K]$, say $S(K+1,N)$, is given by

$$S(K+1,N) = \binom{K+N}{N} = \frac{(K+N)!}{K!\ N!}.$$

Proof. See Pakes (1991) Lemma 32.

Note that Lemma 2 implies that the number of grid points that need to be evaluated grows as a polynomial in N. The approximation methods we introduce presently are designed to reduce the number of grid points that need to be evaluated at each iteration. In particular we will present an approximation technique for which the number of points that need to be evaluated at each iteration of the fixed point calculation is independent of the number of firms ever active in the industry.

Still, at each grid point we do evaluate, we need to evaluate the value function at every achievable industry structure in the following period. For example, recall that in our case the outcome of the firm's research endeavors could lead the firm's ν_1 to either equal 0 or 1, while the increment in value of the outside alternative could have the same realizations. So if n is the number of incumbents who remain active, there are 2^{n+1} industry structures to evaluate with no entry, and the same number with entry. Lemma 2 can be used to reduce the number of distinct industry structures that need to be evaluated, but we will still be left with a number of structures that grows as a Kth order polynomial

in n. We indicate below how one can adapt a technique due to Kortum(forthcoming) to reduce the computational burden at each grid point from being polynomial in n to being linear in n.

We note that the combination of approximation techniques we shall introduce also make the computational burden of the fixed point calculation independent of the fineness of the grid (of K in our notation). As a result they could be used to approximate the value function for problems defined on continuous state spaces (as well as for the discrete problems discussed here).

As for the third component of the calculation of the required CPU time, the number of iterations needed for convergence, we expect it to, at least eventually, become independent of N. As we increase N, the effect of an additional active firm on the value of being at a particular point ought to diminish, so the final iteration for the value function calculated at the N−1 firm equilibrium should be closer to the N−firm value function we are looking for.

We calculated 2 to 6 firm equilibria (i.e. N=2,...,6) for different values of the parameter vector for our example, and found that the computational time for the 6 firm equilibria was about 5 hours on our sparc 1. The no−entry barrier equilibria value of N for most of these runs was 6 firms. However the time required to calculate the equilibria went up by a factor of 5 as we went from an N to an N+1 equilibria. This in spite of the fact that the number of iterations required before our convergence criteria was met typically got multiplied by fractions between .5 and .7 when we moved up N in units of one (though this was more variant between runs). Thus, though the computational techniques discussed thus far may suffice for computing equilibria for markets with a small number of agents, we will need to improve on them in order to analyze many of the markets of interest.

One possible source of improvement is to consider computational techniques that attempt to reduce the computational burden of obtaining the equilibrium by fitting the value function at only a small fraction of the points in S, and then using the information

obtained from those values to predict the value function at other points as needed. More generally, all we require is an approximation to a function which determines optimal policies at any point in S, and there are many different ways of doing this. The symposium in the JEBS (1990) reviews and compares several different approximating techniques in the context of computing equilibria for a representative agent stochastic growth model. Judd (1990), sketches a general framework and computes equilibria from models with two agents (no entry or exit), and Marcet (1990) reviews progress in this field to date.

Many of these techniques fit polynomials in a set of functions that span, or form a basis for, a "rich enough" collection of approximating functions (the Chebyshev or Legendre polynomials for example) to a small set of points, and then use the fitted polynomial to predict the other points as needed. We begin by showing how to embed such approximations into the computational algorithm described in section II.

The heterogeneous agent problems we are interested in are by nature multidimensional, the dimensionality of the state vector for any given agent going up with the number of other agents active in the market.[7] So our polynomial approximating functions will be maps from Ω^N into $\mathbb{R}$. A function f which takes Ω^N into $\mathbb{R}$ is a polynomial of order λ if for all $\omega \in \Omega^N$

$$(3) \quad f(\omega_1,...,\omega_N) = \sum_{p=0}^{\lambda} \sum_{h_N=0}^{p-\Sigma h_i} \cdots \sum_{h_1=0}^{p} \alpha(h_1,...,h_N)\omega_1^{h_1}...\omega_N^{h_N}$$

$$\equiv \sum_{h \in H^N} \alpha(h)\omega(h)$$

where $h = \{h_1,...,h_N\} \in H^N$, $H^N = \{h \in \mathcal{Z}^N_+ : \Sigma h_i \leq \lambda\}$, and $\alpha(h) \in \mathbb{R}^N$ for all $h \in H^N$.

---

[7]Throughout we will consider the case where N, the least upper bound to the number of agents ever active in a given period, is less than or equal to $K = \#\Omega$. In this case the dimensionality of the state vector is smaller when we calculate value functions as a function of the vector of ω values of all active agents. When $N \geq K$, use of a counting measure on Ω as the state vector minimizes the dimensionality of the state space.

38

The collection of all such polynomials (obtained by varying α), together with the usual operations of addition and scalar multiplication, is a vector space (over the real numbers), say $\mathcal{V}_\lambda$. A basis for this vector space is the set of tensor products of the $\omega_i^{h(i)}$ with h varied over H^N (see Hoffman and Kunze,1972,section 5.6). These are just the functions implicit in the $\omega(h)$ in equation (3). Though we do not pursue it here, the following discussion could be generalized by looking for an approximation in a vector space spanned by the tensor products of in $g(\omega_i)$ for suitably chosen $g(\cdot)$.

The iterative procedure used to calculate the fixed point defining the value function for our problem can be modified to find an approximating polynomial, a $\hat{V} \in \mathcal{V}_\lambda$, as follows. Define a set of basis points, say $\omega(j) \in \Omega^N$, for j=1,2,.... If there are J basis functions, the basis points must generate at least J linearly independent values for those functions. Starting at some initial guess for the vector α, let the estimate of the coefficients at the k–1th iteration of the recursive calculation be α^{k-1}. Now calculate the value function at the basis points by substituting

$$(4) \quad \hat{V}^{k-1}[\omega(j)] = \omega(j)'\alpha^{k-1},$$

into the kth iteration of the recursive calculation in (II.1a). (II.1b) then produces an x^k, which when substituted back into the (II.1a) that used $\hat{V}^{k-1}$ produces a new value function, say $V*^k(\cdot)$, at each of the basis points. We choose α^k to minimize the Euclidean distance between $\omega'\alpha^k$ and $V*^k(\cdot)$ at the basis points. That is, if W is the matrix formed from the rows $\omega(j)$,

$$(5) \quad \alpha^k = [W'W]^{-1}W'V*^k.$$

This procedure can be generalized slightly by approximating a monotone function of $V(\cdot)$ by a polynomial in the basis functions, instead of approximating $V(\cdot)$ itself.

Without further restrictions the number of functions needed to form a basis for $\mathcal{V}_\lambda$, and hence the minimum number of points at which we need to fit the value function for this approximation, still grows as a *polynomial in* N. However, we have not yet used the fact that the value function is exchangeable in the vector $(\omega_2,...,\omega_N)$. If we restrict our search to the subspace of $\mathcal{V}_\lambda$ that satisfy the restriction that, for all $\omega^N \epsilon \Omega^N$,

$$(6) \quad \hat{V}(\omega_1;\omega_2,...,\omega_N) = \hat{V}(\omega_1;\pi_2,...,\pi_N),$$

for any N–1 dimensional vector $\tau = (\pi_2,...,\pi_N)$ which is a permutation of $(\omega_2,...,\omega_N)$, we reduce the number of required basis functions dramatically. Indeed, provided N$\geq\lambda$, the number of required basis functions becomes *independent of* N. That is the content of the following theorem.

7. Theorem

The space of polynomials of order λ satisfying equation 6, together with the usual operations of addition and scalar multiplication, is a vector space, say $\mathcal{V}_\lambda^e \subseteq \mathcal{V}_\lambda$, with dimension

$$(8) \quad \dim \mathcal{V}_\lambda^e \leq \Sigma^N_{p=0}\Sigma^p_{i=0}\Delta(i)=\varphi(\lambda),$$

where $\Delta(i)$ is the number of partitions of the number i (see below). Further, (8) holds with equality if N$\geq\lambda$. Note that $\varphi(\lambda)$ is independent of N.

Proof

The fact that addition and scalar multiplication preserves partial exchangeability proves that the subspace of functions satisfying (6) is a vector space. The proof of (8) is a result of the following lemma.

9. Lemma (proved in Appendix 1 of Pakes, 1991).

An $f \in \mathcal{V}_\lambda$ is also a member of $\mathcal{V}_\lambda^{\mathscr{E}}$, if and only if for all $h \in H^N$,

$$\alpha_f(h_1; h_2, \ldots, h_N) = \alpha_f(h_1; \pi_2, \ldots, \pi_N),$$

for any $(\pi_2, \ldots, \pi_N)$ which is a permutation of $(h_2, \ldots, h_N)$. ∘

Define $m_j(h)$ to be the jth largest element in the vector $(h_2, \ldots, h_N)$ for $j=1, \ldots, N-1$ (using any tiebreaking rule that preserves the natural order of pairs that are ordered). Then the lemma implies that we can form a basis for $\mathcal{V}_\lambda^{\mathscr{E}}$ by simply adding together the basis functions from $\mathcal{V}_\lambda$ that have

$$\alpha_f(h_1; h_2, \ldots, h_N) = \alpha_f(h_1; m_2, \ldots, m_N),$$

for each distinct value of the vector $(h_1, m_2, \ldots, m_N)$. What remains is to determine the number of distinct α coefficients this generates. Let $p(h)$ be the order of the basis function corresponding to $\alpha(h)$, that is $p(h) = \Sigma h(i)$. Then the number of distinct α coefficients generated by h vectors with $p(h) = p$, and a particular value of h_1, is the number of ways the number $p-h_1$ can be allocated among $N-1$ locations (without regard to order). If $N \geq \lambda \geq p-h$, this is simply the number of partitions of $p-h_1$, or $\Delta(p-h_1)$ (see below). Consequently, the number of distinct α coefficients required to generate all distinct coefficients for the pth order basis functions is $\Psi(p)$, where

$$\Psi(p) = \sum_{i=0}^{p} \Delta(i)$$

$\varphi(\lambda)$ is derived by summing this equation over $p=0,1,\ldots,\lambda$. ∘

Note that the theorem applies to polynomial approximations to any function which satisfies the exchangeability restriction in (6), and need not be used solely for approximating value functions emanating from dynamic stochastic equilibria with heterogeneous agents. For example, Berry, Levinsohn, and Pakes (1991) show how to use Theorem 9 to reduce the dimensionality of polynomial approximations to optimal instrument vectors in estimation problems emanating from Nash Equilibria (see Chamberlain,1990, for the development of the optimal instrument formula, and Newey,1990, for the use of standard polynomial approximations to them; the example used in the Berry, Levinsohn and Pakes paper is estimation of a generalized version of the differentiated product model discussed in Section I of this paper).

We now provide a set of basis functions for $\mathcal{Y}_\lambda^{\mathscr{C}}$. The two distinct first order coefficients implied by the theorem are

$$\alpha(1;0,...,0), \quad \text{and} \quad \alpha(0;1,0,...,0).$$

Consequently the basis functions associated with these coefficients are

$$\omega_1, \quad \text{and} \quad \Sigma_{i=2} \, \omega_i.$$

The four distinct second order coefficients are

$$\alpha(2;0,...), \ \alpha(1;1,0,...), \ \alpha(0;1,1,...), \ \text{and} \ \alpha(0;2,0,...)$$

with basis functions

$$\omega_1^2, \ \omega_1\Sigma_{i=2} \, \omega_i, \ \Sigma_{i_1=2}\Sigma_{i_2=2}\omega_{i_1}\omega_{i_2} \ \text{and} \ \Sigma_{i=2}\omega_i^2.$$

More generally, the $\Delta(p-j)$ p^{th} order coefficients with $h_1 = j$ are

$$\alpha(j;p-j,0,...), \quad \alpha(j;p-j-1,1,0,...), \quad \alpha(j;p-j-2,2,0,0,...),$$

$$\alpha(j;p-j-2,1,1,0,...),..., \quad \alpha(j;1,1,...,1,0,...),$$

with associated basis functions

$$\omega_1^j \sum_{i=2}^{N} \omega_i^{h-j}, \quad \omega_1^j \sum_{i_1=2}^{N} \sum_{i_2=2}^{N} \omega_{i_1}^{h-j-1} \omega_{i_2}, \quad \omega_1^j \sum_{i_1=2}^{N} \sum_{i_2=2}^{N} \omega_{i_1}^{h-j-2} \omega_{i_2}^2,$$

$$\omega_1^j \sum_{i_1=2} \sum_{i_2=2} \sum_{i_3=2} \omega_{i_1}^{p-j-2} \omega_{i_2} \omega_{i_3},..., \quad \omega_1^j \Sigma_{i_1}...\Sigma_{i_{p-j}} \omega_{i_1},...,\omega_{i_{p-j}}$$

The general formula for $\Delta(q)$ requires fairly detailed notation (see, for eg., Abramowitz and Stegum, 1972, p. 825; it is a sum of Stirling numbers of the second kind). For convenience, we provide a listing of $\Delta(q)$ and $\varphi(q)$ for $q=1,...,12$, in Table 7.

Recall that if a λ-order polynomial is a good approximation to the value function, then we need only calculate the value function at $\varphi(\lambda)$ points. Table 7 indicates that $\varphi(12)=854$. For comparison, the pointwise technique used to calculate the results reported earlier required calculating the value function at 639,000 points; and this for a vector of parameters that generated an ergodic distribution of industry structures with an upper bound of six active firms. Thus, at least for industries with a moderate number of firms, polynomial approximations restricted to the subspace of exchangeable polynomials should allow us to cut the number of points at which we evaluate the value function by several orders of magnitude.

q	$\Delta(q)$	$\varphi(q)$
0	1	1
1	2	3
2	3	7
3	4	14
4	5	26
5	7	45
6	11	75
7	15	120
8	21	186
9	30	276
10	41	407
11	55	593
12	75	854

Table 7
Number of Basis Functions for q^{th} Order Exchangeable Polynomials

A numerical test of the quality of the approximation from the exchangeable basis can be obtained by simply fitting the calculated value function for one or more of our examples to the exchangeable basis, and then considering how "good" a fit one obtains. We now provide a brief summary of our results from these tests to date. It works out that just how well the approximation does depends on both the metric we use (that is on what we are most interested in fitting), and on details of how the approximation is done, so a full analysis of our results will have to await a separate paper.

Pakes (1991) reports on the fit of the value function per se from such an approximation. The results were encouraging for two reasons. First the measure of fit obtained from the exchangeable basis for a given order of approximation was, as the theorem suggests it should be, pretty much independent of the number of firms active in the market. Second, the results were not noticeably different when, instead of fitting the coefficients by projecting the calculated value function at all grid points onto the exchangeable basis, we fit the polynomial to a small number of points first (the points that were mod 3 in the vector sense) and then used the resulting coefficient estimates to predict the value function on the entire set of grid points.

Our primary interest is not in reproducing the value functions per se, but rather in reproducing the policies that those value functions imply. We have not yet fit the polynomial approximating technique into the recursive fixed point algorithm, so we have no clear answer to how well the approximating technique does in this regard. We have, however, found the policies generated by the approximated value function obtained from projecting the actual value function onto the exchangeable polynomial basis. The recursive algorithm would then use these policies as input into a second step approximation to the value function, and then iterate on this procedure. The iterative procedure is likely to generate an internal consistency between the errors in the approximating functions and the errors in the policies that will be absent in our one step procedure.

The results from the one step procedure were better for investment, than for entry

and exit. Largely this is because the investment policy is a smooth functional of the value function, whereas entry and exit policies are, by their very nature, discrete responses. Of course some averaging over time goes on for the exit and entry policies, so that if entry is missed in one period, it may well be compensated for in the next, but at least in some respects the impacts of these errors did not seem to "average out" over the 10,000 simulated periods. For example the approximated policies for the base case discussed in the last section had about 15% of the equilibria points with 5 or more firms active, while the actual base case had 5 or more firms active in only 3% of the cases. On the other hand the approximations generated simulations that were almost exactly on the mark for total investment, the average price—cost ratio, the average of the share of the largest firm, and the welfare measures. So if it were these latter statistics that one were interested in, the approximation error would have been acceptable. We note again that one might do better on the entry and exit policies by fitting the polynomial approximation technique directly into the recursive calculation, and that there are many details of the approximation technique that can be varied in order to try to improve its performance, so these results should be treated as preliminary.

Recall that the CPU time required to compute the value function is a product of; the number of points evaluated at each iteration, the time per point evaluated, and the number of iterations required before convergence. The approximation techniques discussed above have shown how to reduce the dimensionality of the number of points we need to evaluate at each iteration. However, without further development they will actually increase the complexity of the computation needed at each point evaluated. The reason is that at each point we evaluate we require the integral of the value function over the states achievable from that point in the next period, and use of the approximation implies that the values of the value function required for the integrand in this computation must be computed as a product of basis functions and polynomial coefficients (instead of just calling them up from memory, which is what is done when we calculate the value function

pointwise).

We now show how to decrease the computational burden at each grid point by adapting a computational suggestion due to Kortum (forthcoming) to heterogeneous agent models. To begin, note that using our approximation the $k-1^{th}$ iteration of the recursive calculation produces a value function which can be written as

$$(10) \quad V^{k-1}(\omega_1;\omega_2,...,\omega_n) = \Sigma_{l=1}^{\varphi(\lambda)} a_l^{k-1} b_l(\underline{\omega}_1;\underline{\omega}_2,...,\underline{\omega}_n)$$

where

$$\underline{\omega}_i' = (\omega_i,\omega_i^2,...\omega_i^\lambda).$$

Our problem is that we need to have an expression for

$$(11) \quad \Sigma_{\omega_1',...\omega_n'} V^{k-1}(\omega_1';\omega_2,...,\omega_n') p(\omega_1'|\omega_1,x_1,\nu) p(\omega_2'|\omega_2,x_2^{k-1},\nu) \cdots p(\omega_n'|\omega_n,x_n^{k-1},\nu) p(\nu).$$

Let

$$(12) \quad \mu^r(\omega_i,x_i,\nu)=\Sigma_\omega \omega^r p(\omega|\omega_i,x_i,\nu), \text{ and } \underline{\mu}(\omega_i,x_i,\nu) =[\mu^1(\omega_i,x_i,\nu),...,\mu^\lambda(\omega_i,x_i,\nu)].$$

Then (10) and the fact that the basis functions are all sums of polynomials in $[\omega_1,\omega_2,...,\omega_n]$ implies that (11) can be written as

$$(13) \quad \Sigma_\nu \Sigma_{l=1}^{\varphi(\lambda)} a_l^{k-1} b_l[\underline{\mu}(\omega_1,x_1,\nu),\underline{\mu}(\omega_2,x_2^{k-1},\nu),...,\underline{\mu}(\omega_n,x_n^{k-1},\nu)]p(\nu).$$

Note that to obtain (13) as a function of x_1 we have to use the moment generating function for the ω' to compute $(n-1)\lambda$ moments. Given those moments we can obtain the Kuhn Tucker solution for $x_1^k(\omega_1,...\omega_n)$ without explicitly calculating the value of $V(\cdot)$ at all

possible future industry structures. So this procedure does away with the need to evaluate each possible future industry structure at each grid point, and replaces it with a need to calculate the $\lambda(n-1)$ moments of the incumbents state variable. As a result the computational burden at each grid point changes from being polynomial to being linear in the number of active firms. If there are a large number of active firms this will make the computational burden at each grid point smaller using the exchangeable basis of polynomials then in the pointwise calculations. Just how much savings one gets depends on how difficult it is to calculate the vectors $\underline{\mu}(w,x,\nu)$, and this is where a judicious choice of functional forms for the family $\mathbb{P}$ might be helpful.

At least two further points on the computational approximation should be mentioned. First note that once we move to the exchangeable basis of polynomials and adapt Kortum's suggestion for replacing the pointwise calculation of possible future values with calculations of the moments of those values, the computational burden of the algorithm becomes *independent* of the fineness of the grid (in our notation, *of K*). That is K is relevant only in as much as it effects the computation of the moments in (13). If we were to choose a $\mathbb{P}$ which generates an analytic expression for $\underline{\mu}(w,x,\nu)$, then K would be entirely irrelevant (at least from a computational point of view we could, for example, allow the underlying state space to be continuous). Second, we have ignored the possibility of using policy iterations at each iterative step of the calculations to decrease the number of iterations needed before convergence. Since, at least for industries which characteristically have a large number of active firms, the major computational burden of our technique is in obtaining the policies at each grid point, iterating on the given policy is likely to be particularly helpful in reducing the computational burden of our algorithm. Moreover use of the moment's in (13), instead of the actual values of the w_i, should also allow us to reduce the computational burden of the policy iterations (see Kortum,forthcoming, for a detailed discussion).

Bibliography

Abramowitz, M. and I. Stegun (1972); Handbook of Mathematical Functions, Dover Publications, New York.

Berry, S. (1991) "Discrete Choice Models of Oligopoly Product Differentiation," mimeo, Yale University.

Berry, S., Levinsohn, J., and A. Pakes (1991); "An Equilibrium Model of the Market for New Cars", mimeo, Yale University.

Caplin, A. and B. Nalebuff (1991); "Aggregation and Imperfect Competition: the Existence of Equilibrium" Econometrica, Vol. 59, pp. 25–59.

Chamberlain, G. (1987); "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions", Journal of Econometrics, pp.305–34.

Davis, S. and J. Haltwinger, (1990); "Gross Job Creation, Gross Job Destruction: Microeconomic Evidence and Macroeconomic Implications," NBER Macroeconomics Annual, 5, 123–168.

Dixit, A. (1989); "Hysteresis, Import Penetration, and Exchange Rate Pass–Through", Quarterly Journal of Economics, Vol.104, pp 165–177.

Ericson, R. and A. Pakes (1989), "An Alternative Theory of Firm and Industry Dynamics," Discussion Paper No. 445, Columbia University.

Hansen, L. and T. Sargent, (1990), "Recursive Linear Models of Dynamic Economies", Mimeo, the Hoover Institution.

Hoppenhayn H. and R. Rogerson (1990); "Labor Turnover and Policy Evaluation In A Model Of Industry Equilibrium", mimeo, Graduate School of Business Stanford University.

Journal of Business and Economic Statistics (1990); "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods", symposium edited by J. Taylor and H. Uhlig, in the Journal of Business and Economic Statistics, pp. 1–55.

Judd, K. (1989); "Minimum Weighted Residual Methods in Dynamic Economic Models",

<u>mimeo</u>, the Hoover Institution.

Judd, K. (1990); "*Cournot Versus Bertrand: A Dynamic Resolution*", <u>mimeo</u>, the Hoover Institution.

Kortum S (forthcoming); "Inventions, R&D, and Industry Growth", Unpublished Doctoral Dissertation, Yale University.

Lambson, V.E. (forthcoming); "Industry Evolution with Sunk Costs and Uncertain Market Conditions", <u>Review of Economic Studies.</u>

Mankiw, N.G., and M. Whinston (1986); "Free Entry and Social Inefficiencies", the <u>Rand Journal of Economics,</u> Vol.17, pp 49–58.

Marcet, A.(forthcoming); "Computation of Dynamic Models", in <u>Advances in Econometrics</u>, Proceedings of the 6th World Congress of the Econometric Society, J.J.Laffont and C. Sims(ed.).

Maskin, E. and J. Tirole (1988a) "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs," <u>Econometrica</u>, Vol. 56, pp. 549–569.

Maskin, E. and J. Tirole (1988b) "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles" <u>Econometrica,</u> Vol. 56, pp. 571–599.

Maskin, E. and J. Tirole (1987) " A Theory of Dynamic Oligopoly, III: Cournot Competition," <u>European Economic Review,</u> Vol. 31, pp. 947–968.

Newey, W. (1990); "Efficient Instrumental Variable Estimation of Nonlinear Models", <u>Econometrica</u>, Vol.58, pp 809–39.

Olley, S. and A. Pakes (1991) "The Dynamics of Productivity in the Telecommunications Equipment Industry", <u>mimeo</u>, Yale University, January 1991.

Pakes, A.(forthcoming); "Dynamic Structural Models: Problems and Prospects. Part II; Mixed Continuous Discrete Models and Market Interactions", in <u>Advances in Econometrics</u>, Proceedings of the 6th World Congress of the

Econometric Society, J.J. Laffont and C. Sims (ed.).

Nelson, Richard R. and Sidney G. Winter, (1982); <u>An</u>

<u>Evolutionary</u> <u>Theory</u> <u>of</u> <u>Economic</u> <u>Change</u>, Belknap Press of

Harvard University Press, Cambridge, MA.

Thomas, L.G. (1990); "Regulation and Firm Size: FDA Impacts on Innovation",

the <u>Rand</u> <u>Journal</u> <u>of</u> <u>Economics</u>, pp 497—518.

Tirole, J. (1989) <u>The</u> <u>Theory</u> <u>of</u> <u>Industrial</u> <u>Organization</u>, MIT Press, Cambridge,

Massachusetts.