Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model

Author(s): Ariel Pakes and Paul McGuire

Source: *The RAND Journal of Economics*, Vol. 25, No. 4 (Winter, 1994), pp. 555-589
Published by: Wiley on behalf of RAND Corporation
Stable URL: https://www.jstor.org/stable/2555975
Accessed: 07-01-2019 16:24 UTC

RAND Journal of Economics
Vol. 25, No. 4, Winter 1994

Computing Markov-perfect Nash equilibria: numerical implications of a dynamic differentiated product model

Ariel Pakes*

and

Paul McGuire**

In this article we develop and illustrate a simple algorithm for computing Markov-perfect Nash equilibria. The advantage of the Markov-perfect framework is that it is flexible enough to reproduce important aspects of reality in a variety of market settings. As a result, we hope that our article and (perhaps improved) versions of the associated algorithms will eventually be a part of a tool kit that allows researchers to go back and forth between the implications of economic theory and the characteristics of alternative datasets.

1. Introduction

■ 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 of the variables we are typically interested in (shares in industry output, investment, productivity, etc.). The nature of these differences will generally have important consequences for the way we analyze issues of interest.

That this is true 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 fact that when firms are differentially situated, their responses will typically be different nonlinear functions of the changing variable (entry and exit are extreme examples of these nonlinearities). As a result, the sum of the individual responses (or the aggregate response) will depend on the detailed characteristics of the distribution of response patterns. Interestingly, the importance of explicitly accounting for heterogeneity in response patterns has been a theme

* Yale University and NBER.

** Yale University and the Economic Growth Center.

Section 2 of this article is a detailed example of a model presented in more generality and in greater technical depth in Ericson and Pakes (forthcoming). More generally, the article owes much to previous work and discussions with Rick Ericson. We are particularly grateful for the detailed comments of Ken Judd, which, together with the comments of Steve Berry, V.V. Chari, an anonymous referee, and the participants in several seminars, have resulted in significant improvements in the current draft. This research was supported by the National Science Foundation through grant nos. SES-9122672 and SES-8821722-01.

of recurrent emphasis in the (recent) empirical literature, which uses micro panel data to study the impacts of policy and environmental changes.[1]

Moreover, often we are after more detail on the distribution of response patterns than just the sum of the individual responses. Indeed, frequently what we would like to know is how that distribution is related to the primitives of the problem. Obvious examples in which this more detailed knowledge is of primary interest are easy to come by in almost all aspects of economics. Analyses of the link between default probabilities and the market for finance capital and of the effects of regulatory changes on market structure are examples that occur repeatedly in the finance and industrial organization literatures; 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, 1990) makes any analysis of the causes or the effects of job turnover in labor markets highly dependent on the detailed characteristics of the equilibrium distribution of responses from dynamic heterogeneous agent models. (For an analysis of this dependence in one particular setting, see Hoppenhayn and Rogerson (1993).)

Once our models acknowledge the fact that agents do differ, and grant that their actions may impact on one another, then the computation of the responses needed to analyze the kinds of issues discussed above can become quite demanding. This article provides an algorithm for computing Markov-perfect Nash equilibrium responses (Maskin and Tirole, 1988a, 1988b) for a class of dynamic heterogeneous-agent models that is rich enough to reproduce many of the features of firm-level datasets. Our intention is to provide a tool to help researchers do descriptive and policy analysis in a setting that allows them to go back and forth between the implications of economic theory and the information in these datasets.[2]

We illustrate by numerically analyzing the equilibria from a differentiated-products version of the Ericson and Pakes (forthcoming) model of industry dynamics. The model focuses on the heterogeneity and idiosyncratic uncertainty induced by the random outcomes of investment (or research and exploration) processes.[3] In it firms invest to develop profit opportunities (improved goods or techniques of production, or larger stocks of fixed inputs). The outcomes of the investment process are uncertain. Positive outcomes lead the firm to states in which it earns more profits. If the outcomes generate lesser increments than those of 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 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 vector that provides the level of development of the other firms currently active in the industry (this vector will be referred to as the industry structure) and the level of development of alternatives outside the industry. The level of development of the outside alternative evolves exogenously. Entry, exit, and investment

[1] For two examples, see Thomas's (1990) analysis of the impact of changes in FDA regulations on the rate of innovation in the pharmaceutical industry and Olley and Pakes's (1991) study of the impact of deregulation on aggregate productivity growth in the telecommunication equipment industry.

[2] For a discussion of some of the issues that would arise in obtaining estimates of the parameters of models such as those computed here, see Pakes (1992) and the literature cited there.

[3] There is more than one source of idiosyncratic uncertainty that firms react to, and different sources are likely to be more relevant to analyzing behavior in different industries. It is analytically easy to incorporate exogenous (say demand or factor cost) uncertainty into the framework presented here, provided the alternative possible realizations of that uncertainty do not reorder the relative efficiencies of the firms in the industry (see, e.g., Dixit (1989)). If, as in Lambson (1991), relative efficiencies are affected by price realizations, then a more substantive modification of our algorithm is required.

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 vector determining the 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. Because 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), our equilibrium is Markov-perfect Nash in investment strategies in the sense of Maskin and Tirole (1987, 1988a, 1988b). Alternatively, using the earlier terminology of the differential (or difference) game literature, the equilibrium is a closed-loop no-memory (or feedback) equilibrium (see Basar and Olsder (1982), Starr and Ho (1969), and the literature cited in these references).

At the heart of this equilibrium is a stochastic process that generates the industry structures the model emits. That process is ergodic, but its more detailed characteristics depend on the precise values of the model's parameters. This fact motivates our development of the algorithm for computing the equilibria.[4] We then use that algorithm to do more detailed analysis of special cases of the model. Here we begin by simply illustrating the kind of output that can be produced through the numerical analysis of dynamic equilibrium models that allow for idiosyncratic uncertainties.

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 demand and cost functions. 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 following: a free-entry Markov-perfect equilibrium; an initial temporary monopoly followed by free entry; free entry with an institutional constraint that limits the market share of the largest firm to be below some fixed fraction of the overall market; free entry with institutionally induced incremental sunk entry costs (a license must be obtained to enter the market); and a "perfect" cartel. The results from these experiments illustrate the potential usefulness of the computational techniques for policy and descriptive analysis, and they also provide some insights into the implications of the institutional arrangements we analyze.

The algorithm we suggest for computing 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. Section 5 of the article 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 all 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 grid points as needed.

The major result of Section 5 is analytic. It shows that provided the value function of a given agent is symmetric in the state vectors of its competitors, then the number of

[4] Hoppenhayn and Rogerson (1993) also consider computation of equilibrium 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).

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. We close this section with a summary of some early results on the fit of the approximation for our problem.

Section 2 begins with a brief description of each of the primitives of the model, starting with the profit function, then the other primitives determining incumbent behavior, and finally those determining the behavior of entrants. For each primitive we outline the general characteristics it must possess and then go over the detailed example used in the numerical analysis. We then give a verbal characterization of optimal policies and the nature of equilibrium (for more detail, see Ericson and Pakes (forthcoming)). In Section 3 we introduce the computational algorithm,[5] in Section 4 we provide the numerical results, and in Section 5 we consider its computational burden and the approximation techniques. Section 6 is a short summary.

2. The model

■ The state variables determining the firm's perception of its opportunities are an index of the firm's own efficiency level (ω) and a vector determining the efficiency levels of its competitors (s). Formally, $(\omega, s) \in \Omega \times S \subseteq \mathcal{Z} \times \mathcal{Z}_+^{\mathcal{Z}}$, where $\mathcal{Z}$ denotes the integers, so ω is integer-valued and $s = [s_i]$, where s_i is the number of firms at efficiency i [for $i \in \mathcal{Z}$]. The profit function maps each possible combination of these states into the real numbers, or $\pi(\cdot, \cdot) : \Omega \times S \to \mathbb{R}$.

This representation of the profit function is a reduced form for the equilibrium in the spot market, and the detailed characteristics of the market that lead to this equilibrium can vary from example to example. This allows our algorithm to be used in a variety of market settings. Assumption 1 (with the verbal description following it) provides the conditions on the profit function used in Ericson and Pakes (forthcoming). Together with the other conditions to be introduced presently, this assumption insures that any dynamic equilibrium generated by the algorithm developed here will have the properties reviewed below.

Assumption 1.

(i) $\pi(\cdot, \cdot)$ is nondecreasing in ω for all $s \in S$, and there exists a complete preorder on S, say $\succsim$, such that $\pi(\cdot, \cdot)$ is nonincreasing in s, ordered by $\succsim$, for all ω;

(ii) for all $s \in S$, $\lim_{(\omega \to +\infty)} \pi(\omega, s) \leq \bar{\pi} < \infty$ and $\lim_{(\omega \to -\infty)} \pi(\omega, s) < (1 - \beta)\phi$, whereas for all ω and all $s \in \hat{S}_n(\omega)$, where $\hat{S}_n(\omega) = \{s \in S \mid \Sigma_{\omega' \geq \omega} s_{\omega'} \geq n\}$,

$$\pi(\omega, s) \leq (1 - \beta)\phi + o(1/n).$$

Part (i) of this assumption states that profits are increasing in the efficiency parameter (ω) and that one can order industry structures (S) such that profits are decreasing in that order. In (ii), ϕ is the scrap or exit value of the firm and β is the discount rate, so $(1 - \beta)\Phi$ is the annuity value of the recoverable assets of the firm should the firm exit. Roughly then, in addition to bounding profits, (ii) insures that (for given s) there is an ω low enough, and (for given ω) there is a level of competition (an s) stiff enough, to force

[5] We have put both Gauss and C code for our algorithm, together with a description of that algorithm (in both PCL and PostScript format), on our computer system (Pakes, Gowrisankaran, and McGuire, 1993). They can be accessed using FTP by any Internet user. To do so, the user should FTP to "econ.yale.edu," use "anonymous" as login, and his or her own username as password. Then the user should change directory to "pub/mrkv-eqm" and copy all needed files. There is a "read.me" file to start you off. We note that the Gauss version is significantly (about one hundred times) slower than the C version, and that even the C version is two to three times as slow as the version of the algorithm used in this article.

profits below what a firm would earn on its assets by investing them in an alternative activity.

Quite a wide range of markets generate profit functions with these characteristics. The example in this article uses a profit function from a differentiated product model of the sort analyzed intensively in the recent industrial organization literature (see Anderson and De Palma (1992) for more detail on the theoretical and Berry, Levinsohn, and Pakes (1993) for more detail on the empirical analysis of such models). Elsewhere we have used our algorithm to analyze a homogeneous product market with constant marginal costs that vary across firms (see Ericson and Pakes, forthcoming) and a homogeneous product market with constant common marginal costs up to a capacity constraint that varies across firms (see Berry and Pakes, 1993).

We now derive the profit function for the differentiated product model used in our examples. Index goods by j, where goods $j = 1, \ldots, N$ are the goods produced by the firms competing in the industry, and good 0 is the outside good (a composite of all other goods). Each consumer purchases at most one good from the industry. The utility consumer r derives from purchasing and consuming good j is given by

$$U_{rj} = v_j - p_j^* + \epsilon_{rj}, \tag{1}$$

where v_j is an index of the quality, and p_j^* is the price of, good j, r [$=1, \ldots, M$] indexes different consumers, and the ϵ_{rj} represent differences among consumers in the value they attach to good j. Consumer r chooses good j if and only if she prefers it over all the alternatives, that is, if for $q = 0, 1, \ldots, N$,

$$\begin{aligned}\epsilon_{rj} - \epsilon_{rq} &\geq [v_q - v_j] + [p_j^* - p_q^*] \\ &= [v_q - v_0] - [v_j - v_0] + [p_j^* - p_0^*] - [p_q^* - p_0^*] \\ &= g[\omega_q] - g[\omega_j] + p_j - p_q, \end{aligned} \tag{2}$$

where $p_q = p_q^* - p_0^*$, $\omega_q = g^{-1}[v_q - v_0]$, and $g(\cdot)$ is increasing, concave, and bounded. In the specification we compute $\omega_q = v_q - v_0$ (so that g^{-1} is the identity) over most possible values of $v_q - v_0$, so it is easiest to think of ω in this way. We introduce $g(\cdot)$ to put an upper bound on profits without having to explicitly introduce the factors generally thought to bound profits in empirical work.[6]

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 (adding a consumer-specific constant to all alternatives does not change behavior, so we have subtracted the utility from the outside alternative from all choices making it our "numeraire"). As a result, when we refer to increases in the quality (ω) or the price (p) of a product we mean increases relative to the quality or the price of the outside alternative, and in our notation the quality of a product will decrease whenever the improvements to it 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, and this in turn will generate a positive correlation in their profits. Because movements in the v_j will tend to generate negative correlations in the profits of the firms 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.

[6] 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 would explicitly incorporate the income constraints of consumers and decreasing marginal utility of income (see, e.g., Berry, Levinsohn, and Pakes, 1993). This would have similar effects on demand patterns but would be computationally more burdensome.

Let the set $C[\omega_j; p, s]$, where s is the vector 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 assuming the ϵ_{rj} to be drawn from independent (both across products and across individuals) extreme value deviates with distribution function $G(\epsilon \leq \epsilon^*) = \exp[-\exp(-\epsilon^*)]$, the probability that a randomly chosen consumer will choose good j is

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

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, a unique Nash equilibrium exists (Caplin and Nalebuff, 1991) and satisfies the vector of first-order conditions

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

for $j = 1, \ldots, N$. Profits are then given by

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

where M is the number (or measure) of consumers in the market, and it is understood that the price and share vectors are calculated from the spot-market equilibrium conditions in (3) and (4).[7]

We now specify the laws of motion for (ω, s). The distribution for $\omega_{t+1} - \omega_t \equiv \tau_{t+1}$, conditional on different amounts invested (x) in product development, is given by the family $\mathbb{P} = \{P(\cdot \mid x), x \in \mathbb{R}_+\}$. Assumption 2 provides the characteristic of $\mathbb{P}$ assumed in Ericson and Pakes (forthcoming).

Assumption 2. The elements of $\mathbb{P}$ are obtained as the distribution of the difference of two nonnegative random variables, ν_1 and ν, representing, respectively, the increments in efficiency resulting from the firm's own investments and increases in the quality of the outside alternative. ν_1 is stochastically increasing in x, has probabilities that are continuous functions of x, and has realizations that are independent across firms. Further, $p(\nu_1 = 0 \mid x = 0) = 1$, and the support of ν_1 is a finite connected set (uniformly over x). ν is independent of x, has realizations that are common across firms, and has finite connected support.

Note that there can be no increment in efficiency from the firm's own investment process when there is no investment, and when there is investment that increment is stochastically increasing (in the first-order dominance sense) in the amount invested. ν is an exogenous random variable that represents the force of the competition from outside of the industry. Because $\tau = \nu_1 - \nu$, the possibility of advances outside the industry both puts positive probability on negative values of τ and induces a correlation across firms in realizations of τ.

[7] The proof that this profit function satisfies part (ii) of Assumption 1 and is monotone increasing in ω is straightforward. If $h(s) = \sum_q \exp[g(\omega_q) - p\{\sigma(\omega_q, s)\}]$, then one can verify that $\pi(\omega, s_1) \geq \pi(\omega, s_2)$ if $h(s_1) \leq h(s_2)$, so $h(\cdot)$ orders S, and part (i) is also satisfied.

The example used in the computations puts

$$\tau = \nu_1 - \nu, \tag{6}$$

where

$$\nu_1 = \begin{cases} 1 & \text{with probability } ax/(1 + ax) \\ 0 & \text{otherwise} \end{cases}$$

and

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

Note that if we make the time period per decision small relative to the time period in the data, several realizations of τ can occur in a period whose length is equal to the time between observations. This allows us to generate rich distributions of observable increments from the simple specification in (6). (In such a case we would want to adjust the other parameters of the problem, e.g., the discount rate, in accordance with the length of a data period.)

To choose optimal investment and exit policies, incumbents also need a perceived distribution for the number and efficiencies of their future competitors. If we let $\hat{s}_{t+1}$ be the vector of competitors' efficiencies, that is,

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

where $e[\omega_{t+1}]$ is a vector that puts one in the ω_{t+1} spot and zero elsewhere, then the perceived distribution of $\hat{s}_{t+1}$ conditional on s_t will be written as

$$q^{\omega}\{\hat{s}_{t+1} \mid s_t\} = \Sigma_{\nu_{t+1}} q^{\omega}\{\hat{s}_{t+1} \mid s_t, \nu_{t+1}\} p\{\nu_{t+1}\},$$

where ν is the realization of the increment in efficiency of the outside alternative. Note that $q^{\omega}\{\cdot \mid \cdot\}$ embodies the incumbent's beliefs about entry and exit.

Ericson and Pakes (forthcoming) make the following assumptions on the formation of beliefs and then show that these assumptions will indeed be satisfied in the equilibrium they derive.

Assumption 3. (i) $m(s)$ firms enter in each period, where $m : S \to \mathcal{Z}^+$. Each entrant pays an amount $x_m^e > \beta\Phi$, which is nondecreasing in m (as above, Φ is the sell-off value of the firm and β is the discount rate). The entry process is completed at the beginning of the second period when each entrant becomes an incumbent at state $\omega^e \in \Omega^e \subset \Omega$ with probability $p^e(\cdot)$. Ω^e is a compact connected set, and for all s, $m(s) \leq \bar{m} < \infty$.

(ii) There exists a regular Markov transition kernel $Q[\cdot \mid \cdot]$: $S \times S \to [0, 1]$ such that the kernels $\{q^{\omega}[\hat{s}' \mid s]\}$ can be derived from Q as the transition probabilities for the competitors of a firm at ω, i.e., as the transition probabilities for $\hat{s} = s - e(\omega)$. The stochastic kernels, Q and q^{ω}, are weakly continuous; i.e., the operators formed by using either to compute expectations takes the space of continuous bounded functions on S into itself. Finally, the set of feasible industry structures is compact.

Given that $q^{\omega}(\cdot \mid s)$ provides the incumbent's perceived distribution of the market structure of its competitors, the Bellman equation for the firm's maximization problem can be written as

$$V(\omega, s) = \max\{\Phi, \sup_{(x \geq 0)}$$

$$[\pi(\omega, s) - cx + \beta \Sigma V(\omega + \tau, \hat{s} + e[\omega + \tau])\, q^{\omega}[\hat{s} \mid s, \nu]\, p(\tau \mid x, \nu)\, p(\nu)]\}, \quad (8)$$

where c is the cost of a unit of investment. If Φ 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.

Moving to entry, we used the following simple entry model for our analysis. Entry is assumed sequential from an unlimited pool. Each entrant pays a (sunk) setup fee of $x_e(m)$, obtains a draw from $p^e(\cdot)$, 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 $x_e(m)$.[8]

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

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

then

$$m_s = \begin{cases} 0 & \text{if} \quad V^e[s, 1] \leq x^e(1), \quad \text{else} \\ \min\{m \in \mathfrak{Z}_+ : x^e(m) \leq V^e[s, m],\ V^e[s, m + 1] < x^e(m + 1)\}. \end{cases} \quad (9)$$

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.

In the numerical analysis we set $x_e(1) = x^e$ and $x_e(2) = \infty$, so the maximum number of entrants in any given decision 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

$$p[\omega_e \mid \nu] = \begin{cases} 1 \text{ for } \omega_e = \omega^*, & \text{if } \nu = 0 \\ 1 \text{ for } \omega_e = \omega^* - 1 & \text{if } \nu = 1, \end{cases} \quad (10)$$

that is, if the outside alternative does not go up during the setup period, the entrant enters at ω^*, otherwise entry is at $\omega^* - 1$.

Several of Ericson and Pakes's (forthcoming) results on optimal policies have an impact on the nature of our computational algorithm. We gather these into our Proposition 1, and explain them thereafter.

Proposition 1. (i) There exist three boundaries in $\mathfrak{Z} \times \mathfrak{Z}\infty$, say $\underline{\omega}(s)$, $\underline{\underline{\omega}}(s)$, and $\bar{\omega}(s)$, with the following properties: (a) $x(\omega, s) = 0$ if $\omega \notin [\underline{\omega}(s), \bar{\omega}(s)] \subset \overline{\mathfrak{Z}}$; (b) $\chi(\omega, s) = 0$ if and only if $\omega \leq \underline{\underline{\omega}}(s)$; and (c) if $\inf_s[\underline{\omega}(s)] = \underline{\omega}$, $\inf_s[\underline{\underline{\omega}}(s)] = \underline{\underline{\omega}}$, and $\sup_s[\bar{\omega}(s)] = \bar{\omega}$, then $\underline{\omega}$, $\underline{\underline{\omega}}$, and $\bar{\omega}$, are all finite.

(ii) There exists an $m^* < \infty$ such that for all $m > m^*$ and all $s \in S$, $V^e(s, m) \leq x^e_m$.

[8] Much more complicated entry models are consistent with our assumptions, but in the absence of empirical information indicating which complications are relevant, we stayed with the model above. We should note, however, that a slightly more general entry model is programmed into some versions of our algorithm. In the more general case, the potential entrant's cost of entry is a random draw from a distribution with support $[x_{el}(m), x_{eu}(m)]$, and this distribution is stochastically increasing in m. The model with random entry costs has certain computational advantages over the model that does not (see Section 3).

(iii) There exists an $N^* < \infty$ such that for all $n \geq N^*$ and all

$$s \in S_n(1) = \{s \in S \mid \Sigma_{\omega' \geq 1} s_{\omega'} \geq 1\}, V^e(1, s) < x_1^e.$$

Proof. See Ericson and Pakes (forthcoming), propositions 1 and 2.

Proposition 1(i)(a) follows from boundedness of the value function, as this implies that for sufficiently large ω the increment in the value function resulting from an increment in ω is very small. Because the return to investment is determined by the increment in the value function generated by higher values of ω, the boundedness of that function ensures that investment will be zero for all ω greater than some $\bar{\omega}(s)$. Because firms cannot improve their quality index without some investment, states above $\bar{\omega}(s)$ 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 conditional on s, and because S is compact there is an overall upper bound. Similarly, the possibility of exit generates a lower bound for the observed ω states, $\underline{\omega}$. Note that this implies that we can, without loss of generality, assume $\omega \in \Omega = \{1, \ldots, \overline{K}\}$.

Proposition 1 (ii) states that there is an upper bound to entry in any period (m^*), whereas 1 (iii) states that there exists an industry size (N^*) such that whenever there are N^* or more incumbents there will be no more entry. It follows that there will never be more than $N^* + m^*$ firms active in a given period, and because each must have an $\omega \in \{1, \ldots, K\}$, the cardinality of S (the number of attainable industry structures) is no greater than $K^{N^*+m^*}$. So there is only a finite number of elements in S.

Ericson and Pakes (forthcoming) also provide a formal definition and proof of the existence of a rational expectations Markov-perfect Nash equilibrium under these assumptions. Here we suffice with a narrative description of that equilibrium. In it, firm behavior (incumbent and entrant) depends on the perceived distributions of industry structures formalized in the transition functions, $q^{\omega}[\cdot \mid \cdot]$. On the other hand, the investment, entry, and exit choices generated by that behavior, together with the known distributions of τ given alternative values of x and $p^e(\cdot)$, generate an objective distribution of industry structures. The model is considered consistent for a $Q[\cdot \mid \cdot]$ if and only if the objective distribution generated by the investment, entry, and exit decisions that result from a perceived distribution of industry structures equal to $Q[\cdot \mid \cdot]$ is in fact $Q[\cdot \mid \cdot]$. Any such $Q[\cdot \mid \cdot]$ (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 K-dimensional vectors of bounded integers, i.e., $S = \{s = [s_1, \ldots, s_k] : \Sigma_j s_j \leq N^* + m^*\}$, with s_j providing the number of firms at $\omega = j$ (for $j = 1, \ldots, K$). 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 $s^t = (s_t, s_{t-1}, \ldots, s_1)$, 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 article also proves that this transition kernel is ergodic. The ergodic result implies that no matter s_0, the initial industry structure, s_t, will, in finite time, wander into one particular subset, say R, of the possible industry structures (S). Once in this set there is no probability of communicating with states outside of 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 S goes to μ^*, the unique probability measure whose support is R and satisfies $\mu^* Q = \mu^*$. Consequently, sample averages of all

random variables that are functions of s will converge to their expected values with respect to the probability distribution given by μ^*.

Note that though $(1/T)\Sigma s_t \rightarrow \mu^*$ (with probability one), s_t itself never settles down. Rather, the structure of the industry is in perpetual flux. Depending on the nature of $Q[\cdot \mid \cdot]$, we may 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 parameters determining the impact of investments. These in turn can be varied with the demand and cost patterns or with the technological opportunities and the institutional structure of the industry one wants to analyze.

What the ergodic theorem tells us is that if we are willing to let limit properties suffice, then we can analyze them, and how they react to different values of our parameters, without specifying initial conditions. It is silent on both the characteristics of the limiting averages and the likelihood of the alternative sample paths that could lead to them. On the other hand, if we could actually compute the transition kernel, $Q[\cdot \mid \cdot]$, we could provide descriptive and welfare analysis of the impact of alternative events on both the ergodic distribution of industry structures and the sequence of industry structures emanating from the actual initial conditions prevalent in an economy. We turn to this computational task now.

3. Computational algorithm

■ This section provides a computational algorithm that allows us to solve for the equilibrium strategies and the resulting stochastic process generating $\{s(t)\}$ for different parameterizations of the model.[9] The next section uses the algorithm developed here to compute and compare the sequence of industry structures that would be generated by varying the parameters of the example detailed above in a way that approximates alternative institutional regimes.

We begin with the problem of computing the optimal policies, i.e., of computing $\{\chi(\omega, s), x(\omega, s), V^e(s, m)\}$. Once we have these policies, the transition kernel $Q[\cdot \mid \cdot]$ can be either derived analytically (see Ericson and Pakes (forthcoming) or estimated from the empirical distribution of simulated sample paths for industry structures when the simulated paths are obtained by simulating the outcomes from use of the optimal policies by all incumbents and potential entrants (see below).[10]

To obtain the equilibrium policies we compute the value of being active for different values of a firm's own, and its competitors', state variables: $V(\cdot, \cdot)$: $\Omega \times S \rightarrow \mathbb{R}$. It is easiest to think of our algorithm as starting with the value function for an environment

[9] We note that the problem of estimating the model's parameters breaks down into subproblems that allow one to obtain estimators of subvectors of those parameters without actually computing the underlying dynamic stochastic equilibrium. See Pakes (1992) for a discussion of feasible estimation strategies for this class of models.

[10] The estimates of the transition probabilities from any single simulated sample path will yield (uniformly) consistent estimates of the transition probabilities from the recurrent class of states.

that limits the number of active firms to N. We then push N up until it no longer constrains the problem.[11]

Starting with $N = 1$, or the monopolist's problem, we look for the value function and the optimal policy when only one firm is allowed to be active. When

$$N = 1,\ S = \{e(\omega);\ \omega \in \Omega\},$$

or the set of possible industry structures is just the set of vectors with one at the ω element and zero elsewhere. Temporarily assume that the boundaries for $\Omega = [\underline{\omega}, \underline{\omega} + 1, \ldots, \bar{\omega}]$ are known, and let $k = \bar{\omega} - \underline{\omega}$. The Bellman equation for the monopoly case is

$$V(\omega) = \max \{\Phi, \sup_{(x \geq 0)} \pi(\omega) - cx + \beta \Sigma_\tau V(\omega + \tau)\, p(\tau \mid x)\}. \tag{11}$$

Equation (11) defines a fixed point to the operator $T : \mathbb{R}^k \to \mathbb{R}^k$, defined pointwise by

$$Tl(\omega) = \max \{\Phi, \sup_{(x \geq 0)} \pi(\omega) - cx + \beta \Sigma_\tau l(\omega + \tau)\, p(\tau \mid x)\}.$$

Standard arguments show that this operator is a contraction mapping with modulus β (the classic reference here is Blackwell (1965)). As a result, there is a unique vector $\{V(\omega)\} \in \mathbb{R}^k$ that satisfies (11) and it can be calculated as follows. Start with $l^0(\omega)$, any bounded function from Ω to $\mathbb{R}$ [e.g., $\pi(\omega)$]. Calculate $l^1 = Tl^0$ pointwise and iterate (with the ith iteration calculated as $l^i = Tl^{i-1}$) until

$$\max_\omega \mid l^i(\omega) - l^{i-1}(\omega) \mid \equiv \|l^i(\omega) - l^{i-1}(\omega)\| \leq \kappa,$$

where $\kappa > 0$ is any desired degree of accuracy. The contraction mapping theorem assures us that this will occur in a finite number, say $i = *$, of iterations and that

$$\|l^*(\omega) - V(\omega)\| \leq \kappa(1 - \beta)^{-1}.$$

We now come back to the problem of setting Ω, the support for ω. As noted in Section 2, provided our regularity conditions are satisfied, the solution to the monopoly problem will generate a couple, say $[\underline{\underline{\omega}}^m, \bar{\bar{\omega}}^m]$, with the property that the monopolist exits if and only if $\omega < \underline{\underline{\omega}}^m$, and the firm does not invest whenever $\omega > \bar{\bar{\omega}}^m$. Because it is easy to compute the fixed point for the monopolist's problem (even for large k), we compute it for a connected subset of $\mathcal{X}$ that is large enough to contain $[\underline{\underline{\omega}}^m, \bar{\bar{\omega}}^m]$ and let the computation determine these bounds.

These bounds to ω are valid only for the monopoly problem (for $N = 1$). However, as noted in Section 2, there exists an $\underline{\underline{\omega}}$ such that $\chi(\omega, s) = 0$ whenever $\omega \leq \underline{\underline{\omega}}$ (regardless of s, or the industry structure). Ericson and Pakes (forthcoming) show that $\underline{\underline{\omega}}^m \leq \underline{\underline{\omega}}$, i.e., if ω is low enough to induce a monopolist to exit, it is low enough to induce exit when more than one firm can be active. Therefore, $\underline{\underline{\omega}}^m$ is a lower bound for Ω in all subsequent calculations.

Also recall that there is an $\bar{\omega}$, such that $x(\omega, s) = 0$ whenever $\omega \geq \bar{\omega}$ (no matter s). Because there is an upper bound to the support of the increments in ω achievable in any given period, say $\bar{\tau}$, and a firm that does not invest cannot have an improvement in its ω, the existence of $\bar{\omega}$ generates an upper bound to Ω ($\bar{\omega} + \bar{\tau} - 1$). We have not formally shown that $\bar{\bar{\omega}}^m \geq \bar{\omega}$, but this has been true in every example we have actually computed,

[11] Of course, computational efficiency often dictates circumventing parts of this procedure in actual problems.

so our algorithm uses $\bar{\omega}^m + \bar{\tau} - 1$ as the upper bound to Ω for all N and then, after solving for the value and policy functions at a given N, checks to see whether they result in zero investment for all $\omega \geq \bar{\omega}^m$. If this condition were not satisfied we would have to recompute the optimal policies with a lengthened Ω.[12]

We now come back to the computational problem when $N = 2$ assuming that $\Omega = [\underline{\omega}^m, \underline{\omega}^m + 1, \ldots, \bar{\omega}^m + \bar{\tau} - 1] \equiv [1, \ldots, k]$. This is the crux of the explanation of our algorithm, and we provide it in three parts. We begin with an informal description of our algorithm, follow by formalizing it, and conclude by comparing our algorithm to the most obvious alternatives.

The Bellman equation for a firm whose $\omega = \omega_1$ and which has a single competitor with an $\omega = \omega_2$, say $V[\omega_1, \omega_2]$, is given by

$$V[\omega_1, \omega_2] = \max \{\Phi, \sup_{(x_1 \geq 0)} \pi(\omega_1, \omega_2) - cx + \beta \Sigma_{\tau_1, \tau_2, \nu} V[\omega_1 + \tau_1 + \nu, \omega_2 + \tau_2 + \nu] p[\tau_1 \mid x_1, \nu] p[\tau_2 \mid x_2, \nu] p(\nu)\}, \quad (12)$$

for $(\omega_1, \omega_2) \in \Omega^2$. Now for each such (ω_1, ω_2) we also have a second Bellman equation, the Bellman equation for the competitor. Because the model assumes that all the primitives of the problem are symmetric, that second Bellman equation is obtained by permuting the arguments in (12); that is, it is calculated from (12) as $V(\omega_2, \omega_1)$. A Nash equilibrium solves for (the exit and investment) policies, and the associated value functions, for this couple of Bellman equations simultaneously.

Our algorithm computes the Nash equilibrium policies and values iteratively. Let $V^i(\cdot)$ and $x^i(\cdot)$ be the value function and optimal policies computed in the ith iteration, and set $V^0(\omega_1, \omega_2) = l*(\omega_1)$: the "converged" value function from the monopolist's problem. The ith iteration starts with the policy and value functions from iteration $i - 1$. In calculating the policies for the ith iteration we distinguish (ω_1, ω_2) points at which $V^{i-1}(\omega_2, \omega_1) > \Phi$ from points at which this condition is not satisfied. The first subset of points have a competitor which, according to the $i - 1$th iteration of the value function, would remain active. Because $N = 2$, there will be no entry at these points. The second subset of points are points at which there may be an entrant.

For the points at which there is an incumbent competitor, substitute $x^{i-1}(\omega_2, \omega_1)$ for x_2 and $V^{i-1}(\cdot)$ for $V(\cdot)$ on the right-hand side of the two-firm Bellman equation in (12), and use the result to evaluate the first firm's future under alternative possible x choices. Then the optimal choice of x, that is, $x^i(\omega_1, \omega_2)$, can be found as the solution to the Kuhn-Tucker conditions

$$x^i\{\Sigma_{\tau_1, \tau_2, \nu} V^{i-1}[\omega_1 + \tau_1 + \nu, \omega_2 + \tau_2 + \nu] p[\tau_2 \mid x^{i-1}(\omega_2, \omega_1), \nu] \{\partial p[\tau_1 \mid x^i(\omega_1, \omega_2), \nu] / \partial x^i\} p(\nu) - c/\beta\} = 0, \; x^i(\omega_1, \omega_2) \geq 0. \quad (13)$$

Using symmetry we obtain the ith interation's value for the competitor's optimal x, i.e., for $x^i(\omega_2, \omega_1)$, by simply permuting the indices in (13) and resolving this equation. This gives us each element of the couple $\{x^i(\omega_1, \omega_2), x^i(\omega_2, \omega_1)\}$ solely as a function of the output from iteration $i - 1$.[13]

[12] We note that it is often possible to reduce the size of Ω as N increases, and this will increase the computational efficiency of the algorithm.

[13] We are implicitly assuming that there is a unique solution to the Kuhn-Tucker equations in (13). As noted in Ericson and Pakes (forthcoming), sufficient conditions for this to occur are that the sign of $p'_x(\tau \mid x)$ is equal to the sign of τ, and the sign of $p''_{xx}(\tau \mid x)$ is the negative of the sign of τ, for all $\tau \neq 0$. This condition is satisfied in the example computed in Section 4. For cases in which there is more than one solution to (13), choose the solution that maximizes the right-hand side of (12) when we substitute $x^{i-1}(\omega_2, \omega_1)$ for x_2 and $V^{i-1}(\cdot)$ for $V(\cdot)$ in that expression.

Note that an alternative solution procedure would be to substitute $x^i(\omega_2, \omega_1)$ for $x^{i-1}(\omega_2, \omega_1)$ on the right-hand side of (13), combine this with the analogous Kuhn-Tucker equation for $x^i(\omega_2, \omega_1)$, and solve this system of Kuhn-Tucker equations simultaneously for the couple $\{x^i(\omega_1, \omega_2), x^i(\omega_2, \omega_1)\}$. This would require us to solve a nonlinear system of equations in the investments of all incumbent competitors at each grid point at each iteration of the algorithm; such a procedure is much more computationally burdensome then the one used here (it almost squared the computational times for our examples). Because we solve (13) for each (ω_1, ω_2) couple separately, we never need to solve for more than a zero to a single equation in a single variable. Moreover, a judicious choice for $\mathbb{P}$ can ease the computational burden involved in solving for the optimal x choice significantly. Indeed, in the example used in the numerical analysis below there is an analytic expression for $x^i(\omega_1, \omega_2)$ as a function of last iterations policies and values, and this decreases significantly the computational burden at each iteration.

We now come back to the subset of (ω_1, ω_2) at which $V^{i-1}(\omega_2, \omega_1) = \Phi$. This is the subset in which there is a possibility of entry. Calculate

$$V^{ei}(1, \omega_1) = \beta \Sigma_{\tau,\omega,\nu} V^{i-1}(\omega, \omega_1 + \tau + \nu) p^e[\omega \mid \nu] p[\tau \mid x^{i-1}(\omega_1, \omega_2), \nu] p(\nu), \quad (14)$$

the $i - 1$th iteration's analogue of the value of entry in (9).

If $V^{ei}(1, \omega_1) \geq x_e(1)$, then the value of entry is greater than the sunk cost incurred at entry, and entry will occur. In this case we replace $p[\tau_2 \mid x^{i-1}(\omega_2, \omega_1), \nu]$ with $p^e[\omega \mid \nu]$ and $\omega_2 + \tau_2 + \nu$ with ω on the right-hand side of (13) and solve the resultant Kuhn-Tucker equations for $x^i(\omega_1, \omega_2)$. If $V^{ei}(1, \omega_1) < x_e(1)$, we replace both $\omega_2 + \tau_2 + \nu$ and $p[\tau_2 \mid x^{i-1}(\omega_2, \omega_1), \nu]$ with 1 on the right-hand side of (13) and then solve for $x^i(\omega_1, \omega_2)$.[14]

We have provided a procedure for solving for $x^i(\omega_1, \omega_2)$ for all $(\omega_1, \omega_2) \in \Omega^2$. $V^i(\cdot)$ can now be calculated as

$$\begin{aligned} V^i[\omega_1, \omega_2] = \max\{\Phi, \pi(\omega_1, \omega_2) - cx^i(\omega_1, \omega_2) \\ + \beta \Sigma_{\tau_1,\tau_2,\nu} V^{i-1}[\omega_1 + \tau_1 + \nu, \omega_2 + \tau_2 + \nu] \\ p[\tau_1 \mid x^i(\omega_1, \omega_2), \nu] p[\tau_2 \mid x^{i-1}(\omega_2,\omega_1), \nu] p(\nu)\}, \end{aligned}$$

where it is understood that if $V^{i-1}(\omega_2, \omega_1) = \Phi$ (so there is no incumbent competitor), we substitute $\pi[\omega_1]$ for $\pi[\omega_1, \omega_2]$ and either ω with probability $p^e(\omega)$ or 1 with probability one (according as $V^{ei}(1, \omega_1)$ is greater than or less than $x_e(1)$) for $\omega_2 + \tau_2 + \nu$ and the probability $p[\tau_2 \mid x_2^i, \nu]$.

We have just described informally how we calculate policies and values for the ith iteration from the policies and values at iteration $i - 1$ for all $(\omega_1, \omega_2) \subset \Omega_2$. We now introduce notation that allows us to turn this description into an iteration of a fixed point calculation.

Let $l(\omega_1, \omega_2) = [l_1(\omega_1, \omega_2), l_2(\omega_1, \omega_2)] \in [\mathbb{R}^+, \mathbb{R}^+]$ and

$$l = (l_1, l_2) \in [(\mathbb{R}^+)^{kxk}, (\mathbb{R}^+)^{kxk}] \equiv L_1 \times L_2 \equiv L.$$

The couple (l_1, l_2) provides k^2 candidate values for investment and the value function,

[14] Note that we have assumed that when there is a possibility of entry, the entry decision is made before the investment decisions of incumbents. We could have had the entry decision made after the investment decisions, and this difference in the timing of moves might result in a different equilibrium.

respectively. Starting with the value for l at the $i-1$th iteration of the algorithm, we compute the ith value by applying the operator $T : L \to L$ defined pointwise by

$$Tl(\omega_1, \omega) = [T_1 l(\omega_1, \omega_2), T_2 l(\omega_1, \omega_2)],$$

where the $T_i : L \to L_i$, for $i = 1, 2$, are defined as follows. let $I\{\cdot\}$ be an indicator function and define

$$I_1(\omega_1, \omega_2) = I\{l_2(\omega_2, \omega_1) - \Phi > 0\}$$

and

$$I_2(\omega_1, \omega_2) = I\{\beta \Sigma_{\tau,\omega,\nu} l_2^{i-1}(\omega, \omega_1 + \tau + \nu) p^e[\omega \mid \nu] p[\tau \mid l_1^{i-1}(\omega_1, \omega_2), \nu] p(\nu) \geq x_e\},$$

so that $I_1(\omega_1, \omega_2)$ equals one if the $i-1$th iteration's values indicate that a competitor at (ω_2, ω_1) would remain active and zero otherwise, and $I_2(\omega_1, \omega_2)$ equals one if an entrant would enter were the competitor to drop out and zero otherwise; let

$$\begin{aligned} T_1 l(\omega_1, \omega_2) = I_1(\omega_1, \omega_2) \operatorname{argmax}_{x \in \mathbb{R}^+} [-cx + \beta \Sigma_{\tau_1,\tau_2,\nu} l_2^{i-1}[\omega_1 + \tau_1 + \nu, \omega_2 + \tau_2 + \nu] \\ p[\tau_1 \mid x, \nu] p[\tau_2 \mid l_1^{i-1}(\omega_2, \omega_1), \nu] p(\nu)\} + [1 - I_1(\omega_1, \omega_2)] \operatorname{argmax}_{x \in \mathbb{R}^+} \\ [-cx + \beta \Sigma_{\tau_1,\tau_2,\nu} l_2^{i-1}[\omega_1 + \tau_1 + \nu, \omega I_2(\omega_1, \omega_2) + \{1 - I_2(\omega_1, \omega_2)\}] \\ p[\tau_1 \mid x, \nu] \{p^e[\omega \mid \nu] I_2(\omega_1, \omega_2) + [1 - I_2(\omega_1, \omega_2)]\} p(\nu)] \end{aligned}$$

whereas

$$\begin{aligned} T_2 l(\omega_1, \omega_2) = I_1(\omega_1, \omega_2) \max [\Phi, \pi(\omega_1, \omega_2) - cT_1 l(\omega_1, \omega_2) + \beta \Sigma_{\tau_1,\tau_2,\nu} l_2^{i-1}[\omega_1 + \tau_1 \\ + \nu, \omega_2 + \tau_2 + \nu] p[\tau_1 \mid T_1 l(\omega_1, \omega_2), \nu] \; p[\tau_2 \mid l_1^{i-1}(\omega_2, \omega_1), \nu] \; p(\nu)\} + [1 - I_1(\omega_1, \omega_2)] \\ \max [\Phi, \pi(\omega_1) - cT_1 l(\omega_1, \omega_2) + \beta \Sigma_{\tau_1,\tau_2,\nu} l_2^{i-1} [\omega_1 + \tau_1 + \nu, \omega I_2(\omega_1, \omega_2) \\ + \{1 - I_2(\omega_1, \omega_2)\}] p[\tau_1 \mid T_1 l(\omega_1, \omega_2), \nu] \{p^e[\omega \mid \nu] I_2(\omega_1, \omega_2) + [1 - I_2(\omega_1, \omega_2)]\} p(\nu)]. \end{aligned}$$

Given l^i we compute $l^{i+1} = Tl^i$, and repeat this computation iteratively until $\|l^i - l^{i-1}\| \leq \kappa$ for some sufficiently small κ, or until some preset upper bound to the number of iterations is reached.[15]

If we do converge to a fixed point, say l^*, then l_1^* and l_2^* provide, respectively, the investment policy and the value function from a Markov-perfect Nash equilibrium when there is a preset upper bound of two to the number of firms ever active. To formally check that l^* is indeed an equilibrium we need only note that any such solution, together with the exit policy $\chi(\omega_1, \omega_2) = 1$ if and only if $l_2^*(\omega_1, \omega_2) > \Phi$ and the Markov transition matrix that the solution generates, satisfies the equilibrium conditions (condition 6) in Ericson and Pakes (forthcoming). More intuitively, it is clear that at l^* each incumbent and potential entrant (i) uses as its perceived distribution of the future states of its competitors the actual distribution of future states of its competitors and (ii) chooses its policies to maximize its expected discounted value of future net cash flow given this perception of the distribution of the future states of its competitors.

[15] It is often useful to keep track of both $\|l_1^i - l_1^{i-1}\|$ and $\|l_2^i - l_2^{i-1}\|$, as the two norms can be useful in isolating causes of problems as they occur.

Having found a Markov-perfect Nash equilibrium with $N = 2$, we now push N up to 3 and do the iterative calculation again, starting at $l^0[\omega_1, \omega_2, \omega_3] = l^*\{\omega_1, \max[\omega_2, \omega_3]\}$, with the obvious change in the dimensionality of l and L. The iterative procedure for $N = 3$ is analogous to the procedure for the $N = 2$ case. However, now either one, both, or none of a firm's competitors may exit and, depending on both the number of competitors exiting and the results of the entry calculation, there may be either zero, one, or two entrants.

On the other hand, there is a further simplification that can be used when $N > 2$. Note that symmetry implies that $V[\omega_1, \omega_2 = \kappa_1, \omega_3 = \kappa_2] = V[\omega_1, \omega_2 = \kappa_2, \omega_3 = \kappa_1]$ for any $(\omega_1, \kappa_1, \kappa_2) \in \Omega^3$; that is, we may reorder the competitors of a firm without affecting either that firm's value or its optimal policy. If we impose this property at each iteration of the algorithm, we need not calculate values and policies for all $(\omega_1, \omega_2, \omega_3) \in \Omega^3$, but rather we can restrict ourselves to the subset $\{(\omega_1, \omega_2, \omega_3) \mid (\omega_1, \omega_2, \omega_3) \in \Omega^3$, and $\omega_2 \geq \omega_3\}$. Analogously, for the N-firm equilibrium we can restrict our calculation to the subset of points in $\{(\omega_1, \omega_2, \ldots, \omega_N) \mid (\omega_1, \omega_2, \ldots, \omega_N) \in \Omega^N$, and $\omega_{i-1} \geq \omega_i$, for $i = 3, \ldots, N\}$. We discuss this symmetry property and its implications for the computational burden of the algorithm extensively in Section 5.

After finding the Markov-perfect Nash equilibria for $N = 3$, we restart this procedure with $N = 4$ and continue restarting until we reach an N where, for all s with $\Sigma s(\omega) = N$, $V^e(s, 1) < x_e$. At this N there are no possible industry structures with N firms active at which an entrant would want to enter. The minimal N that satisfies this condition (and such an N will exist provided the assumptions of Section 2 are satisfied) is an upper bound to the number of firms that will ever be in the industry (provided, of course, that the initial s has no more than N firms).[16]

We now provide a brief discussion of possible convergence and uniqueness problems with our algorithm. We begin with convergence and note that there is no guarantee that the algorithm just described does converge. In fact, we found two problems that would periodically cause iterations to cycle back and forth between two or more points (values of l).

First, there were cases in which the points in a cycle differed only in whether there was entry at one or more values of the state vector (in the simplest case, entry at iteration i induced a policy function from the entrant's competitors at i that caused the potential entrant to stay out at $i + 1$, which in turn induced the policy function that caused entry at $i + 2$, and so on). This reflects the fact that entry induces a discontinuity in the function determining the incumbent's value of the future (that function jumps when the last iteration's value of entry passes an entry threshold). To circumvent the convergence problem caused by the entry discontinuity, we wrote a version of the algorithm that allowed entry costs to be random. In it the potential entrant's cost of entry is a random draw from a continuous distribution, say $P^{ce}(\cdot)$, with support $[x_{el}, x_{eu}]$, where $x_{el}(1) > \beta\Phi$. If the value of entry at a particular point is κ, then this version of the algorithm assumes incumbents perceive themselves to be facing a distribution of future competitors with no entrant with probability $1 - P^{ce}(\kappa)$, and they perceive themselves to be facing a distribution with an entrant with probability $P^{ce}(\kappa)$. This made the incumbent's policies a smooth function of the last iteration's value of entry and seemed to do away with cycling problems that could be attributed to entry.[17]

[16] In our computations we often stopped at an N at which there were still a few points with $\Sigma s(\omega) = N$ at which there was entry, and substituted the policies from this calculation into the output program which determined the ergodic class of points (see below). We found that if the ergodic class did not contain points with $\Sigma s(\omega) = N$ at which there was entry (or contained only a small number of points at which this was true), then we could stop at this N without distorting the policies too much for our purposes.

[17] Note that we have implicitly assumed here that, as in our numerical example, there is a maximum of one entrant per computational period. An analogous procedure is still feasible, though somewhat more computationally demanding, when there is a possibility of more than one entrant per period.

The second problem occurred only rarely but is related to the nonuniqueness issue. Nonuniqueness can often be caused by situations in which one firm out of two (or several firms out of many) must exit, but either exiting would generate an equilibrium. A related phenomenon can cause cycling in our algorithm. That is, starting with an iteration in which both firms exit, the next-iteration policies for the two firms are derived from a situation in which both perceive that their competitor will exit. This induces each to stay in, which in turn induces both to exit in the following iteration, and so on.

We also wrote a version of our algorithm with a feature designed to circumvent this problem, but unlike the version of our algorithm with random entry costs, it is significantly more computationally burdensome than the original. In this version of our algorithm we assume there is an order to the exit decisions. The firms higher in the order make their exit decision *knowing* the current iteration's exit decisions by firms lower in the order. Our ordering has firms with lower values of ω making their exit decision first. Thus, in the three-firm case with $\omega_3 \leq \omega_2 \leq \omega_1$, if at the $i-1$th iteration $l^{i-1}[\omega_3, \omega_1, \omega_2] = \Phi$, the firm with $\omega = \omega_2$ makes its ith iteration's exit decision based on whether $l^{i-1}[\omega_2, \omega_1, 1] > \Phi$ (recall that a 1 in the third slot indicates a value when there is no third competitor). Note that to order all exit decisions and calculate policies in this manner, we must also provide an order to two firms with the *same* value of $(\omega, \hat{s})$ and then calculate different policies for them. That is, this version of the program must distinguish among the $s(\omega)$ firms at the same $(\omega, \hat{s})$. This increases the dimension of the state space of the computation (we need to keep track of both the $(\omega, \hat{s})$ of the firm and the "order" of the firm among the $s(\omega)$ firms at a particular ω, at least if $s(\omega) > 1$), and hence the computational burden of the algorithm. Still, in the cases where cycling seemed to be caused by the type of scenario described above, using this more complicated version of our algorithm did tend to stop the cycling problem.[18]

Finally, we note that even after ordering exit there may be more than one Nash equilibrium to our problem. As a result, even if we do converge to a solution there is no guarantee that it is the only solution. We have computed several of our examples (including those discussed here) from different initial conditions, and we have always converged to the same fixed point, so nonuniqueness does not seem to be a problem with the simple functional forms we are currently using.

Having described our algorithm and some of the problems with it, we now briefly compare it to some of the more obvious alternative algorithms. We begin with a natural extension of the computational algorithm for the single-firm problem to problems with many firms. Again we start with the two-firm problem and the Bellman equation in (12). The alternative algorithm is also iterative. Start with any feasible policy function and use it to construct a distribution for a firm's competitor's τ_2 conditional on the alternative possible values of (ω_1, ω_2) say $\{p^0[\tau_2 \mid (\omega_1, \omega_2), \nu]\}$. Now substitute these distributions into (12) for $\{p[\tau_2 \mid x_2, \nu]\}$. This produces what would be the first agent's Bellman equation were that agent to hold beliefs about its competitors' behavior that resulted in $\{p^0[\tau_2 \mid (\omega_1, \omega_2), \nu]\}$. Standard arguments show that this Bellman equation is the fixed point of a contraction mapping. That fixed point can be computed in precisely the same way as we computed the fixed point to the contraction mapping for the single-firm problem (see above; any bounded guess for the initial value of $V(\cdot)$, say $V^0(\omega_1, \omega_2) = \pi(\omega_1, \omega_2)$ or $V^0(\omega_1, \omega_2) = V(\omega_1)$, will do).

[18] There are more traditional ways of modifying iterative algorithms to aid convergence, and we tried some of them with a fair bit of success early on in our work. "Dampening" procedures seemed particularly useful. These keep in storage a weighted average of the value functions of the last several iterations, and use this weighted average to form policies. Later in our research, when we thought that we could isolate the causes of the cycles that led to nonconvergence when we found a convergence problem, we switched to algorithms which attacked those causes directly whenever they occurred. We did this even though the dampening procedure, which is less computationally demanding than the procedures we suggest in the last two paragraphs, often did quite well.

This contraction mapping produces a policy and a discounted future value for each incumbent and potential entrant at every value of (ω_1, ω_2). Consequently, it allows us to form the transition probabilities that those policies imply, say $\{p^1[\tau_2 \mid (\omega_1, \omega_2), \nu]\}$. Nothing, however, assures us that $p^1[\tau_2 \mid (\omega_1, \omega_2), \nu] = p^0[\tau_2 \mid (\omega_1, \omega_2), \nu]$ for every $(\omega_1, \omega_2) \in \Omega^2$. Thus the first agent's beliefs about what its competitors would do will not, in general, turn out to be equal to what those competitors would actually do given the chance to choose their own policies (which implies that we are not at a Markov-perfect Nash equilibrium). As a result, we update the perceived distribution of the competitors' policies to $\{p^1[\tau_2 \mid (\omega_1, \omega_2), \nu]\}$, substitute this distribution into the Bellman equation in (12) for $\{p[\tau_2 \mid x_2, \nu]\}$ and compute the optimal policies and values from this contraction mapping. The algorithm continues iterating in this fashion until a suitable convergence criterion has been met.

Note that this algorithm has a fixed point calculation nested inside another fixed point calculation; i.e., to evaluate each iteration of the outer fixed point (the fixed point that iterates on the family $\{p^i[\tau_2 \mid (\omega_2, \omega_1), \nu]\}$), we need to compute an inner fixed point (the fixed point that computes the value function and policies conditional on $\{p^i[\tau_2 \mid (\omega_2, \omega_1), \nu]\}$). Moreover, the inner fixed point is a fixed point of the same dimension as the outer fixed point. This algorithm is thus much more computationally burdensome than our algorithm (its computational burden per iteration is just the burden of solving a single equation in a single variable many times).

We conclude with two possible extensions to our algorithm. These are both designed to decrease the number of iterations our algorithm needs before convergence. The first is to use an analog of what is generally referred to as "policy iteration" in the literature that deals with the computation of solutions to single-agent dynamic problems (see the review in Rust (1992)). To do this, we would hold the policies constant at the values calculated at a given iteration and iterate on the value function in (12) until that function converges for the given set of policies. In terms of our previous notation we would hold $l_1 = l_1^i$ and apply the operator T_2 repeatedly until l_2 converges for the given value of l_1^i (it can be shown that this modified T_2 operator is a contraction mapping, so the policy iteration step is guaranteed to converge). We would then use the "converged" value function to compute the $i + 1$th iteration policies (as discussed above). Our experience with policy iteration in computing the pointwise solutions discussed in this section is that, at least in the examples computed to date, it was not helpful (it invariably increased the processing time needed for the computation).

A second possibility is to iterate on the policies at each iteration holding the value function constant, a procedure we call "value iteration." That is, after calculating the ith iteration policies we would substitute $x^i(\omega_2, \omega_1)$ for $x^{i-1}(\omega_2, \omega_1)$ on the right-hand side of (12) and, still using $V^{i-1}(\cdot)$ for $V(\cdot)$, calculate new policies. Formally, this step holds $l_2 = l_2^i$ and iterates on the operator T_1 until l_1 converges (or until a preset upper bound to the number of iterations is reached; note that the operator defined by this modification to T_1 is not a contraction mapping). Alternatively, one could use any other nonlinear equation solver to obtain a simultaneous solution for the policies of all agents generated by $V^{i-1}(\cdot)$ (for a discussion of the possibilities see Judd, forthcoming). For our examples, the simple types of value iteration we tried did not increase the speed of convergence of the algorithm. We note, however, that it is easy to add either value or policy iteration steps to the iterative algorithm developed here, and they may be more effective when applied to different problems.[19]

[19] In Section 5 we consider the computational burden of our algorithm and note that for problems with a large number of firms we will have to give up on techniques based on pointwise fixed point calculations and move to techniques based on some form of approximation. In related work, discussed briefly in that section, we have developed and tried techniques based on polynomial approximations, and our limited experience with them indicates that both value and policy iteration do lead to improvements in performance when these techniques are used.

4. Numerical illustrations

■ The base case. 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 equilibrium resulting from the differentiated product profit function in (5), the transition probability function in (6), and the following set of parameter values:

$$\begin{aligned}
\delta \text{ (the probability that the outside alternative moves up)} &= .7;\\
\beta \text{ (the discount rate)} &= .925;\\
x_e \text{ (sunk entry cost)} &= .2;\\
\Phi \text{ (scrap value)} &= .1;\\
M \text{ (size of market)} &= 5;\\
a \text{ (determines the efficacy of investment in generating increases in efficiency)} &= 3,\\
mc \text{ (marginal cost of production)} &= 5,\\
e^{g(\omega)} &= \begin{cases} e^{\omega} & \text{if } \omega \le \omega^*,\\ e^{\omega^*}[2 - e^{-(\omega-\omega^*)}] & \text{otherwise.}\end{cases}
\end{aligned}$$

A few comments on these parameter values will help to set the stage for the numerical analysis. The total current cost of producing the output of this industry works out to be about 25, so the sunk entry cost, i.e., x_e, is about 1/125th of the cost of production in a given period. That is, it is not very costly to enter in this, our base case. Note that 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, ...). Finally the functional form chosen for $g(\omega)$, in particular the difference between $g(\omega)$ and ω for sufficiently large ω, ensures both that profits are bounded, and that they are a "smooth" function of ω.

We began by computing the policies and value function of a monopolist facing this environment. Normalizing so that the monopolist's exit state is one, we found $\bar{\omega}$, the highest value of ω at which there is still investment, to be 21. We then set $p^e(\omega = 3) = \delta$, so that a new entrant will enter at $\omega = 3$ with probability δ and at $\omega = 4$ with probability $1 - \delta$, and ran the algorithm described in Section 3. Note that we are assuming that entry costs are a known constant (x_e) in these runs.

The columns labelled "MP" in the tables to follow are a result of substituting the polices computed in the algorithm into an output program. The output program used these policies and a pseudo random number generator to generate 10,000 periods of industry evolution from an initial condition of an industry with one firm at $\omega = 4$, and then computed an assortment of statistics designed to describe the output from this run. The other columns in these tables were constructed in a similar way to the MP column except that they used policies that were computed with a change in either a behavioral assumption or a parameter value. We shall provide more details on these columns below, but for now suffice it to note the following: in the column labelled "$\sigma^* = .65$" we imposed an institutionally created upper bound to the market share of the largest firm equal to .65; 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 from the base case but assumed there was a perfect cartel controlling all pricing, investment, and entry and exit decisions; and in the column labelled "PP" we used the base case parameters but assumed there was a benevolent social planner controlling all pricing, investment, and entry and exit decisions.

TABLE 1 **Characteristics of Ergodic Distribution**
$\delta = .7$, $a = 3$, $\beta = .925$, $x_e = .2$, $\Phi = .1$, $M = 5$, $mc = 5$

	10,000				
Number of time periods:	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	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

Notes: MP = Markov-perfect Nash equilibrium; $\sigma^* = .65$, MP with market share constrained to be below .65; $x_e = 2$, MP with sunk entry costs increased to 2; PP = planner's problem; Coll. = perfect cartel.

Table 1 provides some statistics that help describe the evolution of market structures that the various solutions generated. We shall focus on the Markov-perfect column first. The top part of the table indicates that the equilibrium process characteristically has either three or four firms active in a given period. The bottom part of the table, however, indicates that there was 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. 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.[20]

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.

About 1,800 firms participated in the industry during the 10,000 periods simulated (Table 2, columns MP). Most, however, remained active only a short number of periods. The modal lifespan was only one period, and the median was just barely two. On the other hand, the distribution of lifespans was extremely skewed, with a mean of 18.4 periods and a standard deviation of 73.2. Firms that did survive the initial period went on to be part of the industry for long periods. Relatedly, both mortality and hazard rates decline markedly over the first seven or eight periods, giving the impression that there was an initial "learning" period (the hazard rates shown in the table are estimated, their standard errors being on the order of .005).

Table 3 provides characteristics of the realized values of the firms that participated. The first point to note is that over 90% of the firms that participated in this industry had

[20] In U.S. manufacturing data, cross-industry analysis shows a positive correlation between entry and exit, but once we condition on a given industry and study its evolution over time we find that the sign (and the magnitude) of the correlation between entry and exit tend to vary by industry (see Dunne, Roberts, and Samuelson (1988) and the literature cited therein).

TABLE 2 **Lifetime Distribution**
$\delta = .7$, $a = 3$, $\beta = .925$, $x_e = .2$, $\Phi = .1$, $M = 5$, $mc = 5$
Number of time periods: 10,000

	Percent			Implied Hazard			Cumulative Percent		
Lifetime	MP	$x_e = 2$	PP	MP	$x_e = 2$	PP	MP	$x_e = 2$	PP
1	48.3	17.5	25.0	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	8.8	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 Ever Active	Mean Lifespan	Median Lifespan	Standard Deviation
MP	1,800	18.4	2	73.2
$x_e = 2$	103	191.5	6	428.0
PP	1,301	19.5	2	93.9

Note: For explanation of the column heads, see the notes to Table 1.

a net loss from their endeavor (generated negative realized values). Most got out early and lost a small amount, but there are those that invest for a while, 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 (recall that the sunk cost of entry was only .2, so this implies a benefit/cost ratio

TABLE 3 **Realized Value Distribution**
$\delta = .7$, $a = 3$, $\beta = .925$, $x_e = .2$, $\Phi = .1$, $mc = 5$, $M = 5$

	Realized Values			Lifetime	Sum of Realized Values		
Obs/Num	MP	$\sigma^* = .65$	$x_e = 2$	MP	MP	$\sigma^* = .65$	$x_e = 2$
1	72.8	45.1	83.0	79	72.8	45.0	83.0
2	52.6	38.9	34.0	247	125.4	84.0	117.0
3	33.1	36.3	32.4	718	158.6	120.3	149.4
4	32.7	35.7	28.4	118	191.3	156.0	177.8
5	29.3	30.2	20.2	102	220.6	186.3	198.0
10	22.8	21.2	15.7	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	1514.0	\
171	−.05	−.07	\	2	1717.5	1514.0	\
1491	−.10	−.7	\	4	1586.5	1336.1	\
1800	−4.08	—	—	15	1282.1	—	—

	Mean	Median	Standard Deviation	# Positive	Mean of Positive	# Negative	Mean of Negative
MP	.69	−.1	4.09	170	10	1630	−.27
$\sigma^* = .65$	.62	−.1	3.74	150	9.5	1465	−.28
$x_e = 2$	2.22	−1.89	11.28	39	14.2	74	−2.5

Note: For explanation of the column heads, see the notes to Table 1.

of 50), 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 surviving 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 Section 3). If we hold the ω's of all competitors constant and compute a section of the value function showing how it changes with changes in the firm's own ω, we find it to be initially convex and then concave in ω (see Ericson and Pakes (forthcoming) for a theoretical discussion of when this must happen). Because investment is, roughly speaking, an increasing function of the slope of the 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 increments in the firm'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 that has developed its ω past its inflection point are very low, and a firm that 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 is 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.

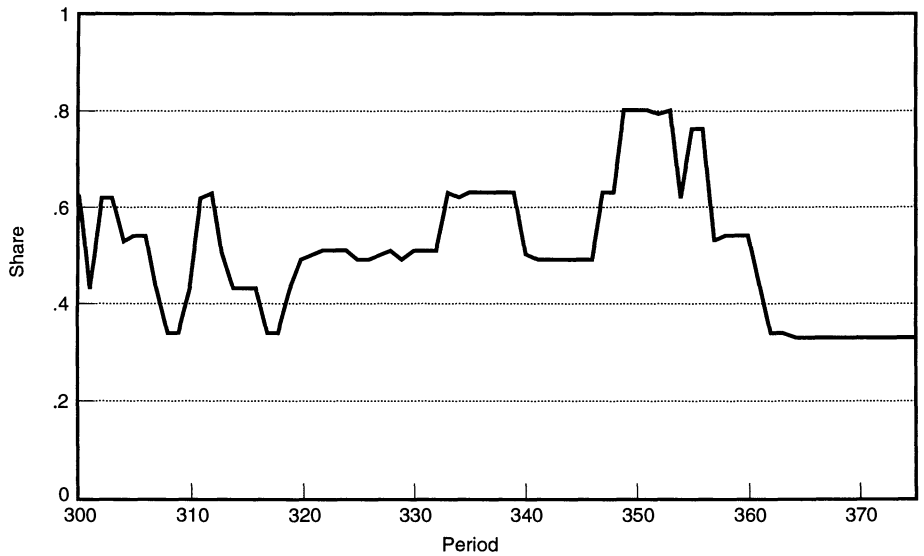
Recall that there is free entry into this industry. As a result, the average of the realized discounted values of new entrants will be quite close to the sunk costs of entry (it will actually be slightly more than that because of integer problems and because of the fact that we have restricted there to be no more than one entrant in a period). If we limit ourselves either to rates of profit or market values of *active* firms, we obtain an indication of supernormal rates of return (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 are dominated by firms that survived for long periods, and *they* indeed earned very high rates of return on their sunk investments. That is, because the observed data are selected on the basis of success, we ought to expect supernormal average returns for active firms even in the most "competitive" of industries.

Tables 4 and 5 (again, in column MP) provide summary statistics on the distribution of the one-firm concentration ratio and on the sales weighted average of the price/cost ratio over the 10,000 periods, whereas Figure 1 plots the evolution of the one-firm concentration ratio over a period. In 90% of the periods, the average markup was between 27% and 44%.

Relatedly, the industry is most often reasonably fractured. The one-firm concentration ratio averaged .39 in an industry in which there are almost always either three or four active firms. Periodically, however, a firm will surge ahead of its competitors and stay there for long periods, creating epochs in which the industry looks quite concentrated. For

FIGURE 1

SHARE OF LARGEST FIRM
$\delta = .7$, $a = 3$, $\beta = .925$, $x_e = .2$, $\Phi = .1$, $M = 5$, $c = 5$



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 ten-period stint at .8. For the next fifty 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.

FIGURE 2

SHARE OF LARGEST FIRM
$\delta = .7$, $a = 3$, $\beta = .925$, $x_e = .2$, $\Phi = .1$, $M = 5$, $c = 5$

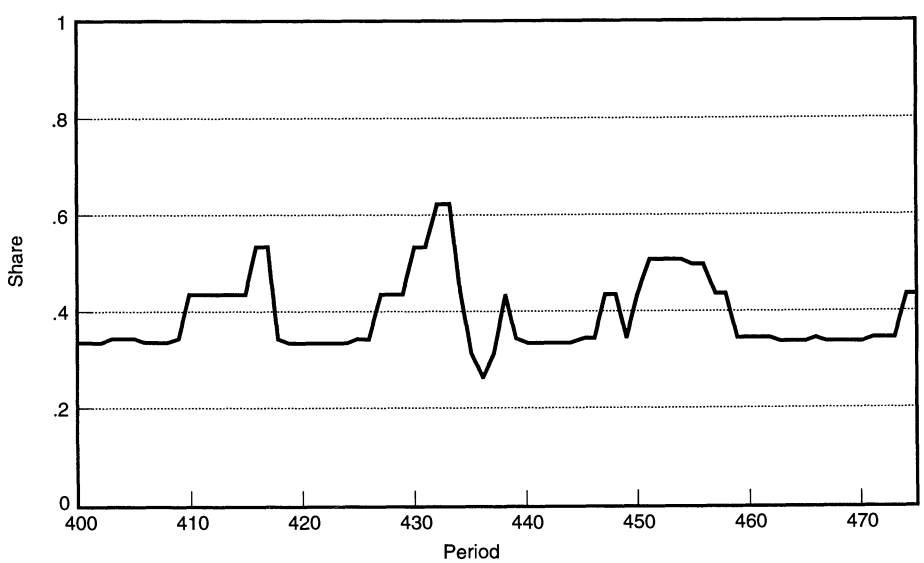


TABLE 4 One-Firm Concentration Ratios

	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

Note: For explanation of the column heads, see the notes to Table 1.

□ **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 welfare. We calculate the welfare associated with each different problem 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 existence in each period. The latter is calculated as

$$\int [\max_j(\omega_j - p_j + \epsilon_j)]dG(\epsilon_1, \ldots, \epsilon_n) = \log \textstyle\sum_j \exp(\omega_j - p_j), \tag{15}$$

where the equality is a result of the fact that $G(\cdot)$ is multivariate extreme value (see McFadden, 1981) 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 (generated by the outcomes of the investment processes and the process determining the value of the outside alternative). As a result, we simulate 100 runs for each

TABLE 5 Price/Cost Ratios

	MP	$\sigma^* = .55$	$x_e = 2$	Coll.
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	1.30	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

Note: For explanation of the column heads, see the notes to Table 1.

institutional environment and then list the resulting means and standard deviations for the discounted net cash flow to the firms, the discounted 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 $\omega = 4$; i.e., $s_0 = e(4)$. 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 that 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.

We use a different (and much simpler) algorithm to calculate the policy functions for the cartel and the social planner. Let $B(\omega_1, \ldots, \omega_n)$ be the one-period (or current) benefits to the decision maker when there are n firms active and their efficiencies are $(\omega_1, \ldots, \omega_n)$. 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 (15) 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

$$\max_{(p_1,\ldots,p_n)} \sum_j M\sigma(\omega_j, p_j; \omega_1, \ldots, \omega_n, p_1, \ldots, p_n)\,[p_j - mc], \tag{16}$$

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

We calculate the value function for the planner and the cartel from the following recursion. If $V^{i-1}(\omega_1, \ldots, \omega_n)$ is the $i-1$th iteration of this function and we have ordered the ω so that $\omega_i \leq \omega_{i-1}$, then we calculate the ith iteration value function as

$$V^i(\omega_1, \ldots, \omega_n) = \max_{(q\in 1,\ldots,n)} \overline{V}^i(\omega_1, \ldots, \omega_q), \tag{17}$$

where

$$\begin{aligned}\overline{V}^i(\omega_1, \ldots, \omega_i) = {} & B(\omega_1, \ldots, \omega_q) + (n-q)\Phi + \max \\ & \{\sup_{(x_1,\ldots,x_q)} - c\textstyle\sum_j x_j + \sum V^{i-1}(\omega_1', \ldots, \omega_n')\, p(\omega_1' \mid \omega_1, x_1, \nu) \ldots p(\omega_q' \mid \omega_q, x_q, \nu)\, p(\nu); \\ & \sup_{(x_1,\ldots,x_q)} - c\textstyle\sum_j x_j - x_e + \sum V^{i-1}(\omega_1', \ldots, \omega_q', \omega_e)\, p(\omega_1' \mid \omega_1, x_1, \nu) \ldots \\ & \qquad p(\omega_q' \mid \omega_q, x_q, \nu)\, p(\omega_e \mid \nu)\, p(\nu)\}.\end{aligned}$$

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

Table 6 provides the welfare results. We begin by comparing the various 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

TABLE 6 Social Welfare from Alternative Market Structures

Benefits/ Market Structure	Total Firm Cash Flows		Consumer Benefits		Total Benefits[a]	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
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[b]	69.8

Note: 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.

[a]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).

[b]The mean of the discounted incremental entry fee is 7.8, when $x_e = 2$.

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. Because 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 be stressed here. First, these institutional arrangements restrict the monopolist to produce no more than one product. Although there may, depending on the source of monopoly power, 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 its members.

The results of the welfare calculations for this institutional environment are given in the row labelled "perfect collusion" in Table 6. As compared to the single-product monopolist, producer surplus, of course, goes up; but the increase is under 5%. Consumers now have a choice of products, however, 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 quite large, large enough to make us worry about case study comparisons of the impacts of different institutional regimes.

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 that 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 5). 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 in which there are typically two or three firms 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 7 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 does the Nash free-entry solution. This is because the entrant and the investor in the Nash game do not take account of the negative effects of their investment and entry decisions on the profits of their competitors, whereas the planner's actions take account of all externalities. The planner does, however, generate an equilibrium with both more firms and more investment than the equilibrium generated by perfect collusion. This is because the planner takes account of the increase in consumer as well as producer surplus generated by entry and increases in product quality, whereas the colluder takes account only of the effect of entry and investment decisions on producer surplus.[21]

Before leaving this comparison we should emphasize that our computations for the "collusive" case simply assumed that there was some outside authority that could ensure that the collusive agreement was maintained (e.g., a legal source of monopoly power). We have not considered whether one could support such a regime with an implicit system

TABLE 7 **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

[21] For a theoretical discussion of the social optimality of the entry rates generated by a free market, see Mankiw and Whinston (1986) and the literature cited therein. Note that Mankiw and Whinston assume that their planner can control entry rates but not (as our planner does) postentry behavior.

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 that 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 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 $\sigma \geq .65$, whereas 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, i.e., 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 5). 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 the effect of the increase in the sunk cost of entry on welfare was not noticeable.

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 1–5 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 periods in the case but in only 1% of the cases after the sunk cost increase), and it 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. 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 were higher in the free-entry Nash than in the 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; 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 that is a salient feature of datasets 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 data and obtaining reasonably reliable estimates of the primitive parameters needed to analyze the nature of equilibrium.

5. Computational burden and approximation techniques

■ We have calculated equilibria for two to six firms (i.e., $N = 2, \ldots, 6$) for a variety of problems. The policies and value functions for a six-firm equilibrium typically take about five hours to compute on our Sun Sparc 1 workstation. Further, the time required to calculate these equilibria usually went up by a factor of about five every time we moved N up by one. These types of calculations indicate that though the computational techniques discussed above may well be adequate for either analyzing industries with a small number of firms, or as a teaching tool, we will need to improve on them before they can be used to analyze larger markets. This section begins with a brief discussion of the magnitude of the computational problem and then introduces approximation techniques designed to reduce the computational burden of the algorithm.

The time required to compute the fixed point that determines optimal policies and the value of those policies is roughly the multiple of (i) the time per grid point evaluation, (ii) the number of grid points evaluated at each iteration, and (iii) the number of iterations until convergence. We begin by considering the number of grid points.

Note first that the number of distinct grid points is k^N, where N is the maximal number of firms ever active and k is the number of distinct elements in Ω. So if there were no further simplifications, the number of distinct grid points grows exponentially in N. Luckily, symmetry[22] implies that

$$V(\omega_1, \omega_2, \ldots, \omega_N) = V(\omega_1, \omega_{\pi(2)}, \ldots, \omega_{\pi(N)}) \tag{18}$$

for any $N - 1$ dimensional vector $\pi = [\pi(2), \ldots, \pi(N)]$ which is a permutation of $(2, \ldots, N)$. It follows that we need not evaluate the value function at each distinct grid point. Indeed, the number of points one needs to evaluate is the number of distinct N-element vectors with $\omega_2 \geq \omega_3, \ldots, \geq \omega_N$ and $1 \leq \omega_i \leq k$, for $i = 1, \ldots, N$. That number can be computed from the following proposition.

[22] In probability theory a function that satisfies (18) would be called "exchangeable" in its last $N - 1$ arguments. We started out using this terminology but were persuaded to change to "symmetric" to conform to terminology used by economic theorists.

Proposition 2. The number of distinct sequences $[\omega_1, \ldots, \omega_N]$, with $\omega_i \geq \omega_{i-1}$ and $\omega_i \in [1, \ldots, k]$, say $S(k, N)$, is given by

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

Proof. See Pakes (1992), Lemma 32.

Proposition 2 implies that the number of grid points that need to be evaluated grows as a (kth-order) polynomial in N.

We have used (18) to reduce the dimension of the fixed-point calculation in the numerical analysis that underlies all the examples we have computed to date. So the computational times given above, times that we argued were too large to make it feasible for us to use our techniques on industries with large numbers of firms, already incorporate the savings that symmetry allows for. As a result, we looked for approximations that reduce the number of grid points even further. Below we show how we can manipulate simple approximation techniques in a way that reduces dramatically the number of grid points that need to be evaluated. Indeed, these techniques make the number of grid points that need to be evaluated at each iteration *independent* of N.

Still, at each grid point we do evaluate, we need to evaluate the value function at every achievable industry structure in the following period. Taking our example, recall that the firm's research endeavors could lead its ν_1 to equal either zero or one, whereas the increment in value of the outside alternative could have the same realizations. So if n is the number of incumbents that remain active, there are 2^{n+1} industry structures to evaluate with no entry, and the same number with entry. Proposition 2 can be used to reduce the number of distinct industry structures that need to be evaluated, but this still leaves us with a number of structures that grows as a kth-order polynomial in n.

In a separate article, Pakes and McGuire (1992) show how one can combine a moment generating function technique similar to that introduced in Kortum (1992), with the approximation techniques introduced below to lead to an algorithm that both does away with the dependence of the number of grid points on N and decreases the computational burden at each grid point. That article integrates those techniques into the algorithm we are about to describe, adds policy and value iteration steps, and then provides some results on both the computational burden and the numerical accuracy of the resulting procedure.[23]

Our goals for this section are more modest. We begin by showing how easy it is to incorporate approximations into the algorithm outlined in Section 3. We then show how the symmetry property in (18) reduces the dimensionality of the fixed-point calculation when the approximations are used. Finally, we conclude with some brief comments on how well the approximating functions did in fitting the value functions from the pointwise calculations given in our example.

This latter step should be regarded as a first step in evaluating the algorithm that uses our approximations. The procedure looks to see if there is a member of the class of approximating functions that provides an adequate fit to the true value function. If such a function does not exist, then a technique that iterates across members of the class of approximating functions looking for a function that provides an adequate approximation to the true function is not likely to be successful.

The approximation techniques we consider attempt to reduce the computational burden of the iterative procedure 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

[23] The combination of approximation techniques used in Pakes and McGuire (1992) also makes the computational burden of the fixed point calculation independent of the fineness of the grid (of K in our notation). As a result, they apply equally to problems with continuous as well as discrete state spaces.

function at other points as needed. More generally, all we require is an approximation to a function that determines optimal policies at any point in S, and there are many different ways of doing this. Taylor and Uhlig (1990) review and compare 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 (and no entry or exit), and Marcet (1992) reviews progress in this field.

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 3.

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. 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 $\omega \in \Omega^N$

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

$$\equiv \sum_{h \in H(N)} \alpha(h)\, \omega(h) \tag{19}$$

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

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 $\mathcal{V}_\lambda$. It can be shown that a basis for this vector space can be formed from the functions $\omega(h)$ with h varied over H^N, as in (19) above (see Hoffman and Kunze (1972)).[24]

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 the vectors $\omega^j \in \Omega^N$, and compute the vector of basis functions for each of those points, say $w(\omega^j)$, where w: $\Omega^N \to \mathbb{R}^J$, for $j = 1, 2, \ldots$. Here J is the number of basis functions, and we require the basis points to generate at least J linearly independent values for those basis functions. Starting at some initial guess for the vector α, let the estimate of the coefficients at the $i-1$th iteration of the recursive calculation be α^{i-1}. These coefficients determine the ith iteration's approximation to the value function at any state, say ω^*, as

$$\hat{V}^{i-1}[\omega^*] = w(\omega^*)'\alpha^{i-1}. \tag{20}$$

Substitute (20) for $V(\cdot)$ on the right-hand side of equations (12), (13), and (14) at each of the basis points. This system can be used to calculate the ith iteration's entry, exit and investment policies in precisely the same way as (12), (13), and (14) were used to update the optimal policies in Section 3. Next, substitute these policies together with $\hat{V}^{i-1}$ into the right-hand side of the Bellman equation in (12). The left-hand side of that equation now gives us a new value function, say $V^{*i}(\cdot)$, at each of the basis points. We choose α^i

[24] We note that the following discussion could be generalized by looking for an approximation in a vector space spanned by the tensor products of $g(\omega_i)$ for suitably chosen $g(\cdot)$ or by looking for an approximation to a monotone function of $V(\cdot)$ instead of an approximation for $V(\cdot)$ itself. This added generality is often helpful when one knows something about the properties of $V(\cdot)$ and can choose transformations that make it easier for the basis functions to mimic those properties.

to minimize the Euclidean distance between $w(\omega)'\alpha^i$ and $V^{*i}(\cdot)$ at the basis points. That is, if W is the matrix formed from the rows $w(\omega^j)$,

$$\alpha^i = [W'W]^{-1}W'V^{*i}. \tag{21}$$

Now iterate on this procedure until convergence is achieved (where we define convergence in terms of closeness of the value functions and of the policies between adjacent iterations).

Have we gained much from using the approximation? Unfortunately, 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 (though this time as a polynomial of order λ rather than of order k). However, we have not yet used the fact that the value function is symmetric in the vector $(\omega_2, \ldots, \omega_N)$. If we restrict our search to the subspace of $\mathcal{V}_\lambda$ that satisfy the restriction that, for all $\omega^N \in \Omega^N$,

$$\hat{V}(\omega_1, \omega_2, \ldots, \omega_N) = \hat{V}(\omega_1, \omega_{\pi(2)}, \ldots, \omega_{\pi(N)}), \tag{22}$$

for any $N-1$ dimensional vector $\pi = (\pi_2, \ldots, \pi_N)$ which is a permutation of $(2, \ldots, N)$, 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 proposition.

Proposition 3. The space of polynomials of order λ satisfying equation (22), together with the usual operations of addition and scalar multiplication, is a vector space, say $\mathcal{V}_\lambda^{\mathcal{E}} \subseteq \mathcal{V}_\lambda$, with dimension

$$\dim \mathcal{V}_\lambda^{\mathcal{E}} \leq \sum_{p=0}^{\lambda} \sum_{i=0}^{p} \Delta(i) \equiv \varphi(\lambda),$$

where $\Delta(i)$ is the number of partitions of the number i (see below). Further, if $N \geq \lambda$, $\dim \mathcal{V}_\lambda^{\mathcal{E}} = \varphi(\lambda)$. Note that $\varphi(\lambda)$ is independent of N.

Proof. The fact that addition and scalar multiplication preserves partial symmetry proves that the subspace of functions satisfying (22) is a vector space. The proof of the proposition is a result of the following lemma.

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

$$\alpha_f(h_1, h_2, \ldots, h_N) = \alpha_f(h_1, h_{\pi(2)}, \ldots, h_{\pi(N)})$$

for any $[\pi(2), \ldots, \pi(N)]$ which is a permutation of $(2, \ldots, N)$.

Proof. (See Appendix 1 of Pakes (1992).)

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 tie-breaking rule that preserves the natural order of pairs that are ordered). Then the lemma implies that we can form a basis for $\mathcal{V}_\lambda^{\mathcal{E}}$ 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). For example, if $p = 5$ and $h_1 = 1$, we look for the number of ways in which we can partition the number 4 (which is 5: (4), (3, 1), (2, 2), (2, 1, 1), and (1, 1, 1, 1)).

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 $p = 0, 1, \ldots, \lambda$. *Q.E.D.*

To use this proposition we need a set of basis functions for $\mathcal{V}_\lambda^{\mathcal{E}}$, and we now outline how to obtain them.[25] The two distinct first-order coefficients implied by the theorem are $\alpha(1; 0, \ldots, 0)$ and $\alpha(0; 1, \ldots, 0)$. Consequently, the basis functions associated with these coefficients are ω_1 and $\sum_{i=2}\omega_i$. The four distinct second-order coefficients are $\alpha(2; 0, \ldots)$, $\alpha(1; 1, 0, \ldots)$, $\alpha(0; 1, 1, \ldots)$, and $\alpha(0; 2, 0, \ldots)$ with basis functions

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

More generally, the $\Delta(p - j)$ pth-order coefficients with $h_1 = j$ are

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

with associated basis functions

$$\omega_1^j \sum_{i=2}^{N} \omega_i^{h-j}, \quad \omega_1^j \sum_{i1,i2=2}^{N} \omega_{i1}^{h-j-1}\omega_{i2}, \quad \omega_1^j \sum_{i1,i2=2}^{N} \omega_{i1}^{h-j-2}\omega_{i2}^2$$

$$\omega_1^j \sum_{i1,i2,i3=2}^{N} \omega_{i1}^{h-j-2}\,\omega_{i2}\,\omega_{i3}, \quad \ldots, \quad \omega_1^j \sum_{i1,i2,\ldots,i(p-j)=2}^{N} \omega_{i1}\omega_{i2} \ldots \omega_{i(p-j)}.$$

The general formula for $\Delta(q)$ requires fairly detailed notation (see, for example, Abramowitz and Stegun (1972)) so for convenience we provide a listing of $\Delta(q)$ and $\varphi(q)$ for $q = 1, \ldots, 12$, in Table 8.

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 8 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, even for industries with a moderate number of firms, polynomial approximations that are restricted to the subspace of symmetric polynomials should allow us to cut the number of points at which we evaluate the value function by several orders of magnitude.

A first step towards a numerical analysis of the quality of the approximation we get from embedding the symmetric basis into our algorithm can be obtained by simply fitting

[25] The same proposition and basis functions apply to polynomial approximations to any function that satisfies the symmetry restriction in (22) and ought, therefore, to be useful in a variety of economic applications. For example, Berry, Levinsohn, and Pakes (1993) use these results to reduce the dimensionality of a polynomial approximation to the optimal instrument formula in their econometric analysis of demand and supply in differentiated product markets.

TABLE 8 **Number of Basis Functions for qth-Order Exchangeable Polynomials**

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

the calculated value function for one or more of our examples to the symmetric basis. We now provide a brief summary of our results from these tests to date.

Pakes (1992) 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 symmetric 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.

The measures of fit we obtained from projecting the symmetric basis onto the true function were extremely good by traditional standards, especially given that we were using only sixth-order approximations. On the other hand, it is not clear how one should interpret these numbers, because our primary interest is not in reproducing the value functions per se but rather in reproducing the policies that those value functions imply. In this context, measures of goodness of fit of the policies generated from our approximation might be more useful than measures of fit from the value function per se.

We can obtain the policies implied by our approximations by substituting the approximation to $V(\cdot)$, say $\hat{V}(\cdot)$, for $V(\cdot)$ in the left-hand side of equations (12), (13), and (14) and computing the policies that this approximation generates. Note, however, that a recursive approximation algorithm would use the policies that are the output of this calculation as input into a second-step approximation to the value function, and 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. As a result, we might expect to get a better fit from the policies generated from the iterative algorithm than we get from the one-step procedure we report here.

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 case we computed had

about 15% of the equilibrium points with five or more firms active, whereas the actual fixed point calculation had five 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 we 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 procedure that can be varied in order to try to improve its performance (including increasing the order of the approximation), so these results should be treated as preliminary. The most important lesson we learned from them is that we will have to be quite careful when approximating policies that are not smooth functionals of the value function. Pakes and McGuire (1992) report on several techniques designed to help the recursive algorithm circumvent this problem.

6. Summary

■ This article developed a simple algorithm for computing Markov-perfect Nash equilibria. The advantage of the Markov-perfect framework is that it is flexible enough to reproduce important aspects of reality in a variety of market settings. As a result, we hope this article and (perhaps improved versions of) the algorithms will eventually be a part of a tool kit that allows researchers to go back and forth between the implications of economic theory and the characteristics of alternative datasets.

Section 3 used the computational algorithm to compute the equilibria of a differentiated product version of the Ericson-Pakes model of industry dynamics. The numerical results were quite detailed and served to illustrate the relationship between observable magnitudes, welfare results, and policy alternatives in a much more realistic setting than could have been investigated using analytic techniques. On the other hand, those results are dependent on the particular range of parameter values analyzed in Section 3, and these were not chosen to mimic the behavior of any particular industry.

Section 5 discusses the computational burden of the algorithm. That discussion indicated that the pointwise calculations that underlie the computational framework discussed in Section 2 will have to be abandoned when analyzing behavior in industries in which there are typically a large number of active firms. As a result, we show how the algorithm can be modified to make use of approximations. We then provide a proposition showing that these modifications make the dimension of the basis (and hence the dimension of the fixed point calculation) needed to fit a polynomial approximation independent of N (and not very large).

References

ABRAMOWITZ, M. AND STEGUN, I.A. *Handbook of Mathematical Functions*. 10th ed. Washington, D.C.: Government Printing Office, 1972.

ANDERSON, S.P. AND DE PALMA, A. *Discrete Choice Theory of Product Differentiation*. Cambridge, Mass.: MIT Press, 1992.

BASAR, T.S. AND OLSDER, G.T. *Dynamic Noncooperative Game Theory*. London: Academic Press, 1982.

BERRY, S. AND PAKES, A. "Some Applications and Limitations of Recent Advances in Empirical Industrial Organization: Merger Analysis." *American Economic Review Papers and Proceedings*, Vol. 83 (1993), pp. 247–252.

———, LEVINSOHN, J., AND PAKES, A. "Automobile Prices in Market Equilibrium." NBER Discussion Paper no. 4264, 1993.

BLACKWELL, D. "Discounted Dynamic Programming." *Annals of Mathematical Statistics*, Vol. 2 (1965), pp. 226–235.

CAPLIN, A. AND NALEBUFF, B. "Aggregation and Imperfect Competition: On the Existence of Equilibrium." *Econometrica*, Vol. 59 (1991), pp. 25–59.

DAVIS, S. AND HALTWINGER, J. "Gross Job Creation, Gross Job Destruction: Microeconomic Evidence and Macroeconomic Implications." In O.J. Blanchard and S. Fischer, eds., *NBER Macroeconomics Annual*, Vol. 5. Cambridge, Mass.: MIT Press, 1990.

DIXIT, A.K. "Hysteresis, Import Penetration, and Exchange Rate Pass-Through." *Quarterly Journal of Economics*, Vol. 104 (1989), pp. 205–228.

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

ERICSON, R. AND PAKES, A. "Markov Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies*, forthcoming.

HANSEN, L. AND SARGENT, T. "Recursive Linear Models of Dynamic Economies." Mimeo, the Hoover Institution, 1990.

HOFFMAN, K. AND KUNZE, R. *Linear Algebra*. 2d ed. Englewood Cliffs, N.J.: Prentice Hall, 1972.

HOPPENHAYN, H. AND ROGERSON, R. "Job Turnover and Policy Evaluation—A General Equilibrium Analysis." *Journal of Political Economy*, Vol. 101 (1993), pp. 915–938.

JUDD, K. "Minimum Weighted Residual Methods in Dynamic Economic Models." Mimeo, the Hoover Institution, 1989.

———. "Cournot Versus Bertrand: A Dynamic Resolution." Mimeo, the Hoover Institution, 1990.

———. "Numerical Methods in Economics." Cambridge, Mass.: MIT Press, forthcoming.

KORTUM, S. "Value Function Approximation in an Estimation Routine." Mimeo, Boston University, 1992.

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

MANKIW, N.G., AND WHINSTON, M.D. "Free Entry and Social Inefficiencies." *RAND Journal of Economics*, Vol. 17 (1986), pp. 49–58.

MARCET, A. "Computation of Dynamic Models." In J.-J. Laffont, ed., *Advances in Econometrics, Proceedings of the 6th World Congress of the Econometric Society*. New York: Cambridge University Press, 1992.

MASKIN, E. AND TIROLE, J. "A Theory of Dynamic Oligopoly, III: Cournot Competition." *European Economic Review*, Vol. 31 (1987), pp. 947–968.

——— AND ———. "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs." *Econometrica*, Vol. 56 (1988a), pp. 549–569.

——— AND ———. "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles." *Econometrica*, Vol. 56 (1988b), pp. 571–599.

MCFADDEN, D. "Economic Models of Probabilistic Choice." In C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge, Mass.: MIT Press, 1981.

OLLEY, S. AND PAKES, A. "The Dynamics of Productivity in the Telecommunications Equipment Industry." Mimeo, Yale University, 1991.

PAKES, A. "Dynamic Structural Models: Problems and Prospects, Part II: Mixed Continuous Discrete Models and Market Interactions." In J.-J. Laffont, ed., *Advances in Econometrics, Proceedings of the 6th World Congress of the Econometric Society*. New York: Cambridge University Press, 1992.

——— AND MCGUIRE, P. "Computing Markov Perfect Nash Equilibria II: Approximations." Mimeo, Yale University, 1992.

———, GOWRISANKARAN, G., AND MCGUIRE, P. "Code for Implementing the Pakes-McGuire Algorithm for Computing Markov Perfect Equilibrium." Mimeo, Yale University 1993.

RUST, J. "Dynamic Structural Models: Problems and Prospects, Part I: Discrete Decision Processes." In J.-J. Laffont, ed., *Advances in Econometrics, Proceedings of the 6th World Congress of the Econometric Society*. New York: Cambridge University Press, 1992.

STARR, A.W. AND HO, Y.C. "Nonzero-Sum Differential Games." *Journal of Optimization Theory and Applications*, Vol. 3 (1969), pp. 184–206.

TAYLOR, J.B. AND UHLIG, H. "Solving Nonlinear Stochastic Growth Models: A Comparison of Alternative Solution Methods." *Journal of Business and Economic Statistics*, symposium issue (1990), pp. 1–55.

THOMAS, L.G. "Regulation and Firm Size: FDA Impacts on Innovation." *RAND Journal of Economics*, Vol. 21 (1990), pp. 497–518.

TIROLE, J. *The Theory of Industrial Organization*. Cambridge, Mass.: MIT Press, 1989.