Finite State Dynamic Games with Asymmetric Information: A Computational Framework

by

Chaim Fershtman and Ariel Pakes

August 2004

Finite State Dynamic Games with Asymmetric Information: A Computational Framework.[*]

Chaim Fershtman and Ariel Pakes

July, 25, 2004.

Abstract

We present a simple algorithm for computing an intuitive notion of MPE for finite state dynamic games with asymmetric information. The algorithm does not require; storage and updating of posterior distributions, explicit integration over possible future states to determine continuation values, or storage and updating of information at all possible points in the state space. It is also easy to program. To illustrate we compute the MPE of a collusive industry in which firms do not know each other's cost positions. Costs evolve with the (privately observed) outcomes of their investment decisions. Costly meetings are called when a firm perceives that its relative cost position has improved. The meetings reveal information and realign profits accodingly. We show that parameters determining information flows can effect market structure and through market structure, producer and consumer surplus.

---

[*] We would like to thank Dmitri Byzalov for a very helpful research assistance and the owners of coffee houses in Boston, Chicago, Florence, London, New York, Positano and Tel Aviv whose hospitality enabled us to write this paper.

This paper develops a relatively simple method for computing the Markov Perfect Equilibria of dynamic games with asymmetric information (see Maskin and Tirole (1992, 2001))[1]. We consider a class of dynamic games in which there are n_t active players in each period, each characterized by a vector of state variables. Some of these state variables are publicly observable while others are private information. State variables evolve over time with the outcome of the players' investment process.[2] In each period players' strategies consist of a set of continuous controls (e.g. investment, output choice, prices etc.) and a set of discrete controls (e.g. sending a signal, entry, exit, etc.). Players' payoff at each period depend on the players' state characteristics at that period and their choice of controls. We let the choice of continuous controls affect the probability distribution of discrete state variables, thus enabling us to keep the state space discrete and computable.

Methodologically we show that in games of this form there are simple sufficient statistics such that if agents maximizes with respect to them, then their actions will be Markov Perfect equilibrium strategies. These sufficient statistics are the expected discounted value of future net cash flows given the possible outcomes of their choice of controls. These discounted values are computed conditional on the firms' information sets, and these, in turn, define the possible states of the game.

If one knew the empirical distribution of outcomes from each state one could compute the needed sufficient statistics directly from that distribution and the primitives of the problem. As a result any agent playing the game who had a history of outcomes at its disposal could calculate these statistics by computing averages. Thus players can determine their optimal behavior from a relatively simple set of calculations; in particular the players never have to bother with computing posterior beliefs for the states of their competitors (at least on the set of states that are visited repeatedly or, more formally, visited "infinitely often").

---

[1] We use the term Markov Perfect Equilibrium even though it was introduced in the literature for games with complete information e.g., Maskin and Tirole (2001). For dynamic games with unobservable actions Maskin and Tirole (1992) introduced the term "Markov Sequential Equilibrium". We, however, restrict our attention to a class of dynamic games with asymmetric information for which the state space is finite. For such games we use the standart definition of MPE but extend the state space to include information relevant as well as payoff relevant variables.

[2] We consider a class of games in which the evolution of a players' state variables are only affected by the players' own actions. In the Industrial Organization literature such games are denoted as capital accumulation games.

What we show is that when this empirical distribution is unknown there is a reinforcement learning (or stochastic approximation) algorithm which enables one to compute the equilibrium by a procedure which only requires updating averages (in particular it never requires one to compute posterior distributions or to integrate out over possible future values)[3]. Thus either the players themselves, or a computer algorithm, could use random outcomes and a set of simple calculations to compute the sufficient statistics which determine equilibrium strategies (again, at least on a recurrent class of points). That is the reinforcement learning algorithm converts the seemingly intractable problem of computing the Perfect Bayesian equilibria into a relatively simple problem of updating averages (so simple that we have computed our examples on a five year old laptop computer).

We only expect the algorithm to provide the correct sufficient statistics as the number of iterations increases without bound. However the paper provides a stopping rule which checks whether the algorithm has produced sufficient statistics within any given accuracy of their equilibrium values, and hence can be used to determine when an adequate approximation has been found[4].

To illustrate our algorithm we use it to compute a MPE of an oligopolistic industry organized as a cartel in which firms can sign binding contracts (or otherwise enforce agreements)[5]. Output is determined cooperatively by the firms and depends on the firms' cost positions, but investments (which affect the firms' cost position but include also exit and entry decisions) are done noncooperatively. We assume that firms know their own costs but do

---

[3] The stochastic approximation literature dates back to the classic paper of Robbins and Monroe,1951, and has been used extensively for calculating solutions to single agent dynamic programming problems (see Bertsekas and Tsitsiklis,1996 and the literature they cite). Pakes and McGuire,2001, show that it has significanat computational advanatages when applied to full information dynamic games, but as we will show the advantages in using it to compute the solution to assymetric information dynamic games are much larger.

[4] An alternative approach would have been to view the AI algorithm itself as the way players learn the statistics needed to choose their policies, and justify the output of the algorithm in that way.

[5] Though this is the example that motivated us, there are many other issues that can be studied with our framework. One, that is closely related, is an industry subject to regulation. Firms invest over time in their capabilities but prices and quantities are determined by a regulatory committee and can be changed only after that committee performs a cost analysis. Either the regulated firms, or the regulator, could initiate the (costly) regulatory review.

not observe the random outcomes of the investment processes of their competitors. At times one, or more, of the firms will perceive it in their interest to call a costly meeting of the cartel members in order to realign the output allocation.

The meetings involve a costly verification process that enables a realignment of quantities based on the firms' new cost positions. Meetings may be initiated by a firm that is no longer satisfied with the market allocation agreed upon in the last meeting; say because its investment activities were particularly successful indicating that it would be likely to be allocated higher profits in a new allocation. Our focus is on the interdependence between the collusive regime and the firms' investment activities. The latter determine market structure, cost positions, and welfare[6]. We then use related models to compare cartels with different allocation rules (including a cost efficient allocation), and compare our asymmetric information model to a full information case in which members of the cartel observe the cost positions of their rivals[7].

The next section describes the details of the game in a general setting and provides a definition of a Markov Perfect Equilibrium to this game. Section 2 provides an algorithm which computes the equilibrium to this game. Section 3 introduces our example, section 4 provides the details of how to compute its equilibrium, and section 5 provides numerical results from computing it and related models.

1 A Finite State Dynamic Game with Asymmetric Information.

We extend the framework in Ericson and Pakes (1995) to allow for asymmetric information.[8] In each period there are n_t potentially active firms, and we

---

[6] In a previous paper (Fershtman and Pakes, 2000) we used numerical analysis to study a complete information collusive industry. One of our conclusions was that collusive behavior may be beneficial also to consumers as it triggers larger investment, more variety and less concentrated market structures.

[7] For numerical analysis of dynamic oligopoly games with complete information see Cheong and Judd (2003), Pakes and McGuire (1994), Markovich (2000), Besanko and Doraszelski (2004) Gowrisankaran (1999).

[8] Since we were motivated by our interest in dynamic oligopolies we will call our players "firms" and their payoffs as "profits". In other applications the players might be better

assume that with probability one $n_t \leq \overline{n} < \infty$ (for every t). Each firm has z characteristics arranged into the vector $\omega \in \Omega^z$ where Ω is a finite subset of $\mathcal{Z}^+$, the positive integers. ω would typically include characteristics of the firm's cost function and of its products. Some of ω's components may be public information, and some may be known only to the firm itself. The publicly known components of ω can change from period to period [9].

Letting i index firms we assume that $\omega_{i,t}$ evolves over time with the outcomes of the firm's investment process, say $\eta_{i,t}$, and an exogenous process that affects the ω's of all firms in a given period, say ν_t, both of which takes values in a subset of $\mathcal{Z}^+$, say in $\Omega(\eta), \Omega(\nu)$ respectively, so

$$\omega_{i,t+1} = F(\omega_{i,t}, \eta_{i,t+1}, \nu_{t+1}), \tag{1}$$

where $F : \Omega^z \times \Omega(\eta)^z \times \Omega(\nu)^z \rightarrow \Omega^z$. The distribution of η is determined by the family

$$P = \{P_\eta(.|x, \omega);\ \ x \in R_+,\ \ \omega \in \Omega^z\}, \tag{2}$$

while the distribution of ν is given exogenously. Here x is a control of the firm and we let $c(x) : \mathcal{R}^+ \rightarrow \mathcal{R}^+$ be the cost of x. Note that, at least in this formulation of our problem, we do not allow the investment of a firm's competitors to effect the evolution of its state variables.

In addition to the continuous controls, i.e. x, the firm also chooses among a set of discrete controls, i.e. it chooses an $m \in [0, 1, \ldots, \overline{m}]^m$. In different models these discrete decisions will take on different meanings including; the sending of signals, entry, exit, the calling of meetings among agents, the launching of new products, and so on. These controls update state variables, say $\xi_t \in \Omega(\xi)$ where $\Omega(\xi)$ is a finite subset of $\mathcal{Z}^+$, that are observed by all

---

thought of as members of teams, or regualtors and regulatees, etc.

[9] We have made several simplifying assumptions which are unnecessary. We could have allowed ω and n to take values in all of $\mathcal{Z}$ but then invoked conditions which insure that we only observe ω values on a finite subset and $n \leq \overline{n}$ (see Ericson and Pakes, 1995). Also we are limiting ourselves to information that is either known only to a single agent, or to all agents; we could have allowed for the possiblity of information known to a subset, but not to all, of the agents.

agents, that is[10]

$$\xi_t = \xi(\xi_{t-1}, m_1, \ldots, m_{n(t)}). \tag{3}$$

The information set of each player at period t is the whole observable history up to that period. We would like to focus on a class of strategies that are "Markovian", but the strategy space must include, in addition to "payoff relevant" variables (in the sense of variables that affect current profits), "informational relevant" variables. These are variables that do not affect payoffs directly but they do provide information on the unobserved variables that determine their competitor's actions and, as a result, their own likely future profits. Each firm's period t's information set will be denoted by $J_{i,t} \in \mathcal{J} \subset Z^l$ where $l \leq \bar{l} < \infty$, and Z is a finite subset of $\mathcal{Z}$.

Note that we only consider games in which the state space is finite dimensional. This precludes general history dependent strategies (since an entire history, by definition, does not repeat itself). As will become clear, for our computational algorithm to make sense, we require a subset of states to be visited infinitely often.

At the beginning of each period there is a realization of $\{\eta, \nu\}$, where η represents the outcome of the previous period's continuous control that are only revealed to the firm who took the actions (our ω), and ν represents an exogenous common shock that has an independent effect on profits (e.g. changes in demand conditions or the prices of factor inputs). Firm's then update their information sets with the updating function $\mathcal{I}(\cdot)$ where

$$J_{i,t+1} = \mathcal{I}(J_{i,t}, J_{-i,t}, \eta_i, \eta_{-i}, \nu_t, \xi_t), \tag{4}$$

where $\mathcal{I} : \mathcal{J} \times .... \to \mathcal{J}$. This updating function assumes that we can represent the evolution of the information sets as a Markov process. The fact that $\mathcal{I}(\cdot)$ is a function of $J_{-i,t}$ indicates that information that is private in one period can be come public in the next period.

Firm's expected profits, as well as their strategies, are functions of their information $J_{i,t} \in \mathcal{J}$, i.e.

$$\pi(J_{i,t}) : \mathcal{J} \to \mathcal{R}^+, \;\; x(J_{i,t}) : \;\; \mathcal{J} \to (\mathcal{R}^+)^z, \;\; \text{and} \;\; m(J_{i,t}) : \mathcal{J} \to [0, \ldots \overline{m}]^m.$$

We assume that firms have a common discount rate β , $0 < \beta < 1$.

---

[10] For simplicity we assume that the ξ are deterministic functions of the controls. A more general model would allow for a stochastic disturbance in this relationship.

1.1 Markov Perfect Equilibrium.

A Markov Perfect Equilibrium is a couple;

- $W(\eta|J,m)$ for each $\eta \in \Omega(\eta), m \in [0,\dots \overline{m}]^m$, and $J \in \mathcal{J}$, which will represent the expected discounted value of future net cash flow conditional on a realization of η and a choice of m, and
- strategies $x(J)$ and $m(J)$,

such that:
C1. Strategies are optimal given $W(\cdot)$, that is, they solve

$$max_m sup_x \left[\sum_\eta W(\eta|J,m)p(\eta|x) - c(x) \right].$$

and

C2. The $W's$ are the expected discounted value of future net cash flows when the players play the equilibrium strategies, i.e. if

$$V(J) = \pi(J) + max_m sup_x \left[\sum_\eta W(\eta|J,m)p(\eta|x) - c(x) \right],$$

then

$$W(\eta|J,m) = \beta E[V(J') - c(m,m_{-i})|\eta,m]$$

where $c(m,m_{-i}) : [0,\dots,\overline{m}]^{mn} \rightarrow \mathcal{R}^+$ is the cost firm i incurs as a result of the discrete strategies chosen. Note that the expectation is over (J', m_{-i}), and is formed from the probability distribution of competitors' actions and future locations induced by the rules of motion and the optimal strategies (obtained from C1). Also it is important to the examples we have in mind to allow the discrete strategies of competitors, the m_{-i}, to impose a cost (see below)[11].

---

[11] We have however assumed that the costs of the discrete controls are separable from those of the continuous control.

Remarks.

There are a number of implications of this definition that we use intensively in computing equilibria.

- First note that were we to know the $\{W(\eta|J_{i,t},m)\}$ we would have enough information to compute equilibrium policies for the i^{th} agent in time period t.
- Second note that the sequence $\{(J_{1,t},\ldots,J_{\overline{n},t})\}_{t=1}^{\infty}$ actually observed is a realization from a finite state Markov chain. It follows that it will eventually wander into a recurrent class of points say $\mathcal{R} \subset \mathcal{J}^{\overline{n}}$, and once in $\mathcal{R}$ will remain in it forever (with probability one; see Freedman,1983). So to analyze equilibrium for subgames starting at a point in $\mathcal{R}$ we need only know equilibrium policies on $\mathcal{R}$.
- Third our definition of Markov Perfect Equilibrium for a dynamic game with asymmetric information is a natural extension to the definition of MPE for dynamic games with complete information to the case which includes both payoff and information relevant variables. Note that the definition does not require beliefs about the actual state of competitors, rather all we require is the beliefs needed to form our continuation values, i.e. the $\{W(\eta|J_{i,t},m)\}$.

We have just defined our equilibrium and we are now going to provide a simple reinforcement learning algorithm that can be used to approximate sufficient statistics for equilibrium play. We should note, however, that an alternative approach would have been to view the algorithm itself as the way players learn the statistics needed to choose their policies, and justify the output of the algorithm in that way. A reader who subscribes to the latter approach may be less interested in the section of the paper on testing (which follows the computational section)[12].

---

[12] On the other hand, there are several other issues that arise were one to take this approach seriously, among them; the question of whether (and how) an agent can learn from the experience of other agents, and how much information an agent gains about its value in a particular state from the agent's experience in related states.

2 An AI algorithm to compute equilibrium.

We present a computational algorithm which never requires us to calculate and retain posterior probabilities. Instead we will compute the $W(\eta|J,m)$ in condition C1 iteratively using techniques analogous to those used in the stochastic approximation (or reinforcement learning) literature (see footnote 3). That is, we will start with an initiation procedure for $W(\eta|J,m)$ for $\{(\eta,m,J) \in \Omega(\eta) \times [1,\dots,\overline{m}]^m \times \mathcal{J}\}$, and update these estimates through an iterative process that never requires the explicit calculation of the expectation in our condition C2. Each iteration will only need our estimates of $\{W(\cdot)\}$ at that iteration and the policies implied by those $\{W(\cdot)\}$.

The stochastic approximation algorithm is iterative. An iteration, say k is defined by couple; its location, say L^k, which defines the information set of all agents, and a set of $\{W^k(\cdot)\}$. The algorithm also has an initiation procedure which calculates initial values of $\{W^0(\eta|J,m)\}$ as needed. Starting at any L^0 the algorithm calculates policies for all agents using $\{W^0(\eta|J,m)\}$. These policies, computer generated draws for the realizations of all random variables which determine outcomes conditional on the policies, and the information updating function or $\mathcal{I}(\cdot)$ in (3), determine the updated location, say L^1.

To complete the iteration the algorithm has to update the $\{W(\cdot)\}$. It only updates the components of $\{W(\cdot)\}$ associated with L^0. The draws on (ν,η_{-i}) enable us to construct a draw on J' conditional on (η,m) as

$$\hat{J}'(\eta,m) \equiv \mathcal{I}(J_i^0, J_{-i}^0, \eta, \eta_{-i}^1, \nu^1, \xi^1(m,m_{-i})).$$

$\hat{J}'(\eta,m)$ is then treated as if it were a random draw from the true distribution of J' given (m,η) and J^0. Since $W(\eta|J^0,m)$ is the expectation of $\beta V(\hat{J}'(\eta,m))$ over the true distribution of these draws, the values of $\beta V(\hat{J}'(\eta,m)|W^0)$, calculated from the equation in C2 after substituting W^0 for the true W, are used to update our estimate of $W(\cdot|J^0,m)$. That is

$$W^1(\eta|J^0,m) = .5\beta V(\hat{J}'(\eta,m)|W^0) + .5W^0(\eta|J^0,m).$$

The process is then continued iteratively from L^1 so that the $W^k(\cdot)$ are constructed as the sample average of the simulated draws taken from each location.[13]

---

[13] Any weighting scheme that satisfies Robbins and Monroe's (1951) conditions could be used instead of the simple average used here; i.e. the sum of the weights of each point visited infinitely often must increase without bound while the sum of the weights squarred must remain bounded.

Since our estimates of the $\{W^k(\cdot)\}$ are formed as sample averages, we expect the estimates from a particular location to be more accurate the more times we visit that location. Moreover since the possible locations are finite, the Markov chain that we generate will eventually wander into a recurrent subset of those locations, say $\mathcal{R}$, and once in that subset, will stay their. All points in this $\mathcal{R}$ will be visited infinitely often. Hence we might expect that both the empirical distribution of J' given (η, m, J) and the estimates of $W^k(\cdot)$ associated with points in $\mathcal{R}$ to eventually converge to the true distribution of J' given (η, m, J) and the true expected discounted values of future net cash flows conditional on (η, m, J).

However we are left with the problem of determining whether a candidate set of $\{W^k(\cdot)\}$ outputted by our algorithm do in fact satisfy the equilibrium conditions on $\mathcal{R}$. After providing details of our algorithm, we show that there is a simple test of conditions which insure that this is in fact true.

2.1 Details of the Algorithm.

Each iteration starts with a location, $L^k = (J_1^k, \ldots, J_{\overline{n}}^k)$, and the objects in memory, say $M^k = \{M^k(J) : J \in \mathcal{J}\}$. The elements of L^k specify the information sets of the potentially active agents (for nonactive players the set is empty), and are mutually consistent (the components which are public information have the same values for all agents). The elements of $M^k(J)$ specify the objects in memory at iteration k for information set J. $M^k(J)$ contains a counter, $h^k(J)$, which keeps track of the number of times we have visited J prior to iteration k, and if $h^k(J) > 0$ it contains:

- $W^k(\eta|J, m)$ for $m \in [0, \ldots, \overline{m}]^m$ and $\eta \in [0, \ldots, \overline{\eta}]^z$.
- $\pi(J)$.

If $h^k(J) = 0$ there is nothing in memory at location J. If we require $\{W(\cdot)'s\}$ at a J at which $h^k(J) = 0$ we have an initiation procedure which sets $W^k(\eta|J_i, m) = U(\eta|J_i, m)$ where $U(\eta|J_i, m) \geq W(\eta|J_i, m)$ (the equilibrium values of $\{W(\cdot)\}$ at that location), and we calculate $\pi(J)$[14].

---

[14] Note that we are assuming that there exist known values of $U(\eta|J_i, m)$ that are greater then their equilibrium values. In the games we have dealt with the "default" here would be to use the expected discounted values from a situation where only one firm was allowed in the market for these $U(\eta|J_i, m)$.

2.1.1 Update of the Location.

Recall that $L^k = (J_1^k, \ldots, J_{\overline{n}}^k)$. For each J_i^k that is nonempty, call up $\{W^k(J_i)\}$ from memory and choose $(x(J_i)^{k+1}, m(J_i)^{k+1})$ to

$$max_m sup_x \left[\sum_\eta W^k(\eta | J_i, m) p(\eta | x) - c(x) - c(m_i, m_{-i}) \right].$$

Given $\{x^{k+1}(J_i)\}$ use the $\mathcal{P}_\eta$ distribution and a computerized pseudo random number generator to draw $\{\eta_i^{k+1}\}$, and draw ν^{k+1}. Now set

$$J_i^{k+1} = \mathcal{I}(J_i^k, J_{-i}^k, \eta_i^{k+1}, \eta_{-i}^{k+1}, \nu^{k+1}, \xi^{k+1}(m_i^{k+1}, m_{-i}^{k+1})).$$

for $i = 1, \ldots, \overline{n}$. Set

$$L^{k+1} = [J_1^{k+1}, \ldots, J_{\overline{n}}^{k+1}].$$

2.1.2 Update of the $W's$.

What we need to do is update $W^k(\eta | J_i, m)$ for; $\eta \in [1, \ldots, \overline{\eta}]^z$, $m \in [0, \ldots, \overline{m}]^m$, and $i = 1, \ldots, \overline{n}$. Recall that

$$\hat{J}_i^{k+1}(m, \eta) = \mathcal{I}(J_i^k, J_{-i}^k, \eta, \eta_{-i}^{k+1}, \nu^{k+1}, \xi^{k+1}(m, m_{-i}^{k+1})),$$

Since we can call up $W^k(\cdot)$ in memory even if $h^k(\cdot) = 0$, we can always update as

$$W^{k+1}(\eta | J_i, m) - W^k(\eta | J_i, m) = \frac{1}{h^k(J_{i,t}) + 1} [V(\hat{J}_i^{k+1}(m, \eta) | W^k) - W^k(\eta | J_i, m)],$$

where $V(\cdot)$ is defined as in condition C2 above but with the W^k replacing the W in that condition.

2.2 Testing For an Equilibrium.

We look for a test for the fixed point implicit in C1 and C2 which is not too computationally burdensome, and which has an interpretation as the extent of approximation error in our estimates. The test criteria will be an $L^2(P)$ norm in the percentage difference between the left and right hand sides of the fixed point conditions which define our equilibrium. That is if $V(J)$ is defined as in the first equation in C2 after substituting the $\{W\}$ outputted by the algorithm for those in the formula, we ask if

$$V(J) = \pi(J) + \beta \sum_{\eta} E\left[V(J') - c(m, m_{-i})|\eta, m\right] p(\eta|x) - c(x) \qquad (5)$$

where it is understood that both (m, x) and the probability distribution underlying the expectation are those outputted by the algorithm[15].

There are two problems with direct implementation of the test. First recall that our algorithm never attempts to get accurate estimates of the $\{W(\cdot)\}$ associated at all $J \in \mathcal{J}$, but just for $J \in \mathcal{R}$. Second the actual computation of the required expectation at any $J \in \mathcal{J}$ is extremely computationally burdensome. To compute it we would need to actually calculate posterior distributions and then use those posteriors to integrate out over all possible future J' given J. The algorithm itself was designed specifically to avoid the problems associated with computing these posteriors.

We overcome the first problem, as did Pakes and McGuire (2001), by focusing on a subset of the $\{W(\cdot)\}$ and their associated policies, a subset that is sufficiently large to enable us to analyze subgames starting from points in $\mathcal{R}$. Sufficient conditions for our algorithms policies to be equilibrium policies at a $J \in \mathcal{R}$ differ with whether J is in the "interior" or on the "boundary" of $\mathcal{R}$. Recall that given optimal behavior any $J \in \mathcal{R}$ only communicates with other points in $\mathcal{R}$. Interior points are defined as points that could only communicate with other points in $\mathcal{R}$ no matter the policy chosen while boundary points could communicate with points outside of $\mathcal{R}$ if feasible but non optimal policies were followed.

If J is in the "interior" of $\mathcal{R}$, then (5) must hold for all (η, m). If J is on the boundary then there are $W(\eta|J, m)$ which are constructed from the expectation over $V(J'|W)$ for $J' \notin \mathcal{R}$; that is J' that will not be visited infinitely often. Our algorithm can not be relied on to obtain accurate estimates of these $W(\eta|J, m)$. However accurate estimates are not necessary; all we require is that the estimates of $W(\eta|J, m)$ we do have are larger than the equilibrium values of $W(\eta|J, m)$. Thus if for example $m \in (0, 1)$, J is a

---

[15] We note that, as is the case for all known algorithms designed to compute equilibria for (nonzero sum) dynamic games, there is no guarantee that our algorithm will converge to equilibrium values and policies, nor do we know for sure that there is a unique equilibrium to converge to.

boundary point, and $m(J) = 0$, then we have to insure that $W(\eta|J, m = 1)$ is greater than the equilibrium value of $W(\cdot)$ at this point. To simplify the analogous condition for the continuous controls assume that the family of distributions $\mathcal{P}_\eta$ have the property that their support is the same for every $x > 0$. Then if J on the boundary of $\mathcal{R}$ and our estimates have $x(J) = 0$, we have to insure that the estimates of $W(\eta|J, m)$ are greater than the equilibrium values for all $\eta > 0$. To do so we assume we have exogenous estimates $\overline{W}(\eta|J, m)$ which are greater then equilibrium $W(\cdot)$ and check to see whether the $W(\cdot)$ outputted by our algorithm at boundary points are larger than the associated $\overline{W}(\cdot)$[16].

We circumvent the computational burden of computing the expectations in (5) by substituting *averages over simulated sample paths* for the integrals defining these expectations and then accounting for the simulation error in that average. I.e. since the simulated paths contain sampling error, the squared difference between the $V(J)$ and the average of the sample paths will consist of the sum of

1. the squared difference between $V(J)$ and the expectation of interest (i.e. of the quantity we wish to measure), and
2. the squared difference between the average of the simulated sample paths and this expectation.

(2) has expectation equal to the variance of the sample average and can be unbiasely estimated by the standard estimate of that quantity (the sampling variance in the simulated paths). The law of large numbers insures that the weighted average of these sampling variances across the different J will converge (almost surely) to the true average of these variances. Thus if we square the $L^2(P)$ norm of differences between the $V(J)$ and the simulated

---

[16] At least in Industrial Organization applications there is always at least one set of values for $\overline{W}(\cdot)$ available, the values generated from a monopoly situation. Of course better (i.e. lower) upper bounds may be available. Provided our initiation values are high enough, precisely how one treats boundary points tends to have little impact on actual computational outcomes. This is because boundary points tend to have very small probabilities in the invariant distribution on $\mathcal{R}$, so standard testing procedures which weight with invariant probabilities, are not very sensitive to the accuracy of our estimates at these points. If one were particularly concerned with estimated values and policies near boundary points, then it would make sense to restart the algorithm many times from those points, see Pakes and McGuire (2001).

averages, subtract the weighted average of the estimated variances, and take the square root of the result we obtain a consistent estimate of the $L^2(P)$ norm of interest *without ever computing* the expectations in (5).

Different tests can be formed by simulating paths of either different lengths (say τ), or with different stopping rules (producing different distributions for a random τ). The value of a simulated sample path of τ periods, say $VS(J_i, \tau)$, is calculated as the discounted value of net cash flows earned between the first and τ^{th} period plus β^{τ} times the value the algorithm assigns to the τ^{th} simulated period (or zero if the firm has exited by then). I.e. if $J_{i,\tau}$ is the location of the i^{th} agent in the τ^{th} simulated period, this "terminating" value is just $\beta^{\tau} V(J_{i,\tau})$ as defined in C2 after setting the $\{W\}$ in the formula equal to those outputted by the algorithm.

More formally at each point in $\mathcal{R}$ we

- simulate sample paths for the industry,
- then for each firm
    - calculate the square of the difference between the mean of its simulated values (of the $VS(J_i, \tau)$) and the value computed directly from W (or $V(J_i)$),
    - subtract the estimate of the sampling variance of the mean of the $VS(J_i, \tau)$, and
    - divide the result by $V(J_i)^2$,
- finally average over the agents active at the starting location.

This produces an unbiased estimate of the squared percentage difference between the average value of the sample paths from a given location and the values implied by the $W's$ outputted by our algorithm. Our test statistic is the square root of a weighted average of these differences over locations, with weights set equal to the fraction of times the locations are visited (so, for a long enough run, the weights will converge to their probabilities in the invariant distribution on $\mathcal{R}$).

If we let $E[VS(J_i, \tau)|W]$ be the expectation of $VS(J_i, \tau)$ when the policies used for the simulated run are generated by W, then as either the number of simulation draws per location or the number of locations grows large our test statistic converges to

$$\left\| \left[\frac{1}{n(l)} \sum_{i=1}^{l} \left(\frac{V(J_{i,l}|W) - E[VS(J_{i,l},\tau)|W]}{V(J_{i,l}|W)} \right)^2 \right]^{1/2} \right\|_{L^2(P(\mathcal{R}))},$$

where $n(l)$ is the number of active agents at location l, $P(\mathcal{R})$ is the invariant measure over the locations in $\mathcal{R}$, and we condition the expectation on W to indicate that it is conditioned on the policies generated by W. If the $\{W\}$ were truly equilibrium policies and the computer's calculations were exact the fixed point condition in equation (5) would insure our test statistic tends to zero. In general the statistic is a consistent estimate of the average percentage difference between the two sides of that fixed point condition. We assume we are "at an equilibrium" when it is sufficiently small[17]. Note that the special case of $\tau \equiv 1$ is of special interest as it is the stochastic analogue of the test traditionally used in iterative procedures to determine whether we have converged to a fixed point.

3 Example: An Incomplete Model of A Cartel In A Changing Environment.

There is a large literature in I.O. on price setting arrangements in oligopolistic markets and their stability over time. The theoretical literature has mainly focussed on the trade-off between short run gains from deviations from a collusive agreements and the costs, or penalty, that can be imposed on deviants. A number of observers have noted that the temptations of short run gains is not the only problem a cartel must deal with. Posner (1976, p.65), for

---

[17] Note that a consistent estimate of the variance of this statistic can be obtained by running this procedure many times and calculating the variance in our statistic over these runs. To use this to produce a traditional one-sided statistical test we would need to; (i) decide what is an acceptable percentage error (if for no other reason then to allow for the imprecision of the computer's calculations) and (ii) decide on the size of the test (the probability of type I error we are willing to accept). The size issue is complicated by the fact that by increasing the number of simulation draws we are free to increase the power of any given alternative to one. I.e. we would need to formulate a tradeoff between size, power, and the number of simulation draws.

example, states that "Among the obstacles to fixing a mutually satisfactory price are the conflicting interests of sellers having different costs".[18]

Empirical studies often identify the reconciliation of differences in demands among member firms, often related to differences in the firms' perceived cost positions, as a source of cartel instability. In studying collusive behavior among Bromine producers between 1885-1914 Levenstein (1997) pointed out that half of the price wars, and in particular the more severe ones, occurred when one of the firms demanded a better allocation of the collusive market shares and the other's disagreed[19].

Disagreements are more likely to be a recurrent feature of a cartel's environment if firms are asymmetric, the extent of the asymmetry changes over time, and their is incomplete information on the current positions of the member firms. There have been a number of studies of collusion among asymmetric firms. For the most part the studies assumed an exogenously source of asymmetry that stayed constant over time and focussed on how these asymmetries impacted collusive possibilities (e.g. Schmalensee (1987), Harrington (1989), and Compte, Jenny and Rey (2002)). Our earlier work (Fershtman and Pakes, 2000) allows for endogenously evolving asymmetries which result from entry, exit, and investment processes. As a result, in addition to studying how asymmetries effect collusive possibilities, it also studies how collusion can effect the extent of asymmetry. Still it assumes that all firms' states are public information and that the source of cartel instability is the short run gains from deviation, rather than a desire for a reallocation of power or profits among cartel members.

In two recent papers Athey and Bagwell (2001) and Athey, Bagwell and Sanchirico (2004) studied optimal collusive behavior in an infinitely repeated Bertrand game with private information. Each firm receives a privately observed i.i.d. cost shock every period. The cartel would like to ensure that production is done each period by the low cost firm, but this is difficult to

---

[18] See also Scherer (1980, p.199) who writes "different sellers are likely to have at least slightly divergent notions about the most advantageous price. Especially with homogenous products, these conflicting views must be reconciled if joint profits are to be held near the potential maximum".

[19] More recently studies of the collusive agreements in Lysine came to a similar conclusion; an agreement could not be reached until the productive capabilities of ADM were verified (De Roos, 2000). Interestingly, like in the Bromine cartel, in these cases the price cuts were announced ahead of time, there was no surprise deviation from the prescribed prices and no attempt to gain short run profits.

achieve when cost positions are private information. The first paper allows firms to communicate cost information and make market share proposals before setting prices. The second paper present the collusive arrangement as a repeated adverse selection model with observable actions (prices).

We consider a cartel whose members can *contract* into quantities which can be renegotiated on demand. That is implicitly we are assuming that there is a punishment which is severe enough to deter unannounced deviations (the fact that we do not model that punishment explicitly is one of two reasons for the phrase "incomplete" in the title to this section). Demanding a renegotiation is costly. Firms' invest to improve their cost position, and the stochastic outcomes of a firm's investments are not observed by the firm's competitors. That is the firms' "types" evolve over time, so regardless of whether there was complete revelation of information in one period, behavior in subsequent periods requires beliefs on the likely types of its competitors in those periods. A firm will demand a renegotiation when it believes that the true physical environment has changed in its favor. That is cost positions have changed to be sufficiently more in its favor for it to incur the cost of the renegotiation. The meeting itself is modelled in a reduced form way assuming only that all firms' cost positions are revealed, and that the new allocations insure that firm's that have improved their relative cost positions increase their profits. This is the other reason for the phrase incomplete in the title, and it can be completed in different ways (one example is that of an industry which can hire outside experts to verify the cost position of each firm). The focus of the applied analysis will be on both the observable implications of, and the determinants of welfare in, markets that operate in this way[20].

3.1 An Endogenously Asymmetric Cartel.

We consider an industry in which there are n_t incumbent firms which differ in $\omega_{i,t}$, a characteristic of the firm's cost function which evolves over time with the outcomes of an investment process. To be in the industry a firm must join a cartel. The cartel has periodic meetings in which the quantities each firm markets, say $q_{i,t}$, are determined. Investment, entry, and exit decisions

---

[20] We do not consider the role of the antitrust authorities in determining how our contracts can be written and/or enforced. In contrast, in a recent paper Harringtom (2002) studies optimal collusive price dynamics when price changes affect the probability of a cartel being detected by an antitrust authority. Then the antitrust policy affects cartels behavior.

are, however, made independently. While in force firms do not deviate from their allocated quantities, but each firm reserves the right to call a meeting in which the current allocation is challenged.

The quantities, together with the firms' state variables determine the profits of each active firm, say $\pi_i(\omega, q) = \pi(\omega_{i,t}, q_{i,t}, q_{-i,t})$.[21] Investments are made to improve the firm's "physical" state, their ω value. This evolves over time with the outcome of the investment process and an industry specific exogenous common shock.

We assume that neither the investment itself, nor the output of the investment activity, are public information. So in a typical period each firm knows its own ω_i but does not know the ω's of the other firms. The game is therefore a dynamic game with asymmetric information.

In each period firms can decide whether to abide by the existent quantity agreement or initiate a renegotiation by calling a costly meetings. In the meetings the cost positions of all firms are revealed and a new quantity agreement is formed. Entry and exit also induce a costly realignment of market shares but possibly at a different cost. Potential entrants who do enter pay a sunk cost of entry and enter at a particular state in the following period. Firms who exit receive the scrap value ϕ.

We assume that firms wish to maximize their expected discounted profits and we let β, $0 < \beta < 1$, be the common discount factor for all firms.

3.2 Details of the Example.

We assume a linear market demand function; i.e. $P = a - bQ$ where P is market price and Q is the total quantity produced by all firms. Marginal costs do not change in quantity, but do vary with the firm's productivity or ω i.e., $mc(\omega_i, q_i) = mc(\omega_i)q_i$. We want the quantity vector agreed to in meetings to reflect the firms relative position in the market. We therefore use Schmalensee's (1987) suggestion, and assume that relative market shares are determined by the noncooperative Nash equilibrium of that period, but total

---

[21] Note that the firm's profit does not depend on $\omega_{-i,t}$. For e.g. in Cournot quantity competition the cost of one firm does not affect the profits of its competitors. The issue here is that profits, conditional on quantities, does not reveal anything about the current value of the competitor's states. This would happen in differentiated product models also, if there were many state variables per firm, and current profits did not depend on the state variable for which there is asymmetric information.

output maximizes the sum of profits conditional on these shares[22]. Thus if we let $q^N(\omega)$ be the Cournot equilibrium output vector, and $s^N(\omega)$ be the Cournot equilibrium market shares, then total output ($Q^c = \sum q_i$) is determined as

$$Q^c(\omega) = argmax_Q \sum_i \pi(\omega_i, s^N(\omega)Q)$$

where, $\pi(\omega_i, s^N(\omega), Q) = (a - bQ)s_i^N(\omega)Q - mc(\omega_i)s_i^N(\omega)Q$. All shares are positive if $a/N > \gamma$, and in this case $Q^c(\omega) = \frac{a - \sum_i s_i^N mc(\omega_i)}{2b}$.[23]

There are two types of meetings depending on the circumstances that trigger them. If a continuing incumbent calls a meeting, or their is entry, then market shares are reallocated and the collusive quantities are determined as above. When this occurs all firms bear the cost of FK plus ΔFK for the firms that called the meeting. Entrants pay FK plus the entry costs.

Meetings also occur when a firm exits, since exit triggers the need to reallocate market shares, but we assume that the meeting has a cost of K to each continuing incumbent with $K \leq FK$. Of course if there is only a single firm active there is no need for coordination and the firm may adjust its output without the costs of a meeting.

Evolution of States.

This is a simplified version of our general setup in which

$$\omega_{i,t+1} = \omega_{i,t} + \eta_{i,t+1} - \nu_{t+1}$$

with both η and ν taking values in $\{0, 1\}$, while

$$p(x_{i,t}, \omega_{i,t}) = \frac{A(\omega_{i,t})x_{i,t}}{1 + A(\omega_{i,t})x_{i,t}}.$$

Note that the distribution of $\eta_{i,t+1}$ is better, in the stochastic dominance sense, the larger is investment, $x_{i,t}$.

---

[22] An alternative is to assume a bargaining solution as in Fershtman and Pakes (2000). The only disadvantage of this is that it is computationally more burdensome.

[23] If $s_i^N(\omega) < 0$, we assume that the firm does not produce and then we recalculate the market shares. Also we note that though it is not necessarily true that our assumptions guarantees that collusive profits are higher than the Nash equilibrium profits, this condition is satisfied for the range of parameters that we use in our analysis.

If an incumbent decides to exit it gets a sell-off value of ϕ dollars, exits in the next period and never reappears again. We let $\chi_{i,t} \in \{0,1\}$ indicate whether a firm exits ($\chi_{i,t} = 0$) or continues ($\chi_{i,t} = 1$). Potential entrants decide whether to enter or not. To enter they must pay a sunk cost of x^e which is uniformly distributed over $[x^e_l, x^e_h]$. The realization of $\widetilde{x}^e$ is observed by the potential entrant prior to the entry decision, but is not observable by other players. An entrant appears in the following period as an incumbent at an $\omega_{i,t+1} = \omega^e - \nu_{t+1}$ where ω^e is given. For simplicity we assume there is at most one entrant in every period, and indicate whether entry occurs by the indicator function $\chi^e = \{0,1\}$, $\chi^e = 1$ indicating entry.

Timing of Decisions.

The timing of decisions are as in our general model. At the beginning of the period there is a realization of the outcome of the investment processes of the last period and realizations of the entry cost and the exogenous shocks. If either; (i) a meeting was called in the previous period, or if (ii) an entry decision was taken during the previous period, or if (iii) one or more firms decide to exit during the previous period, then there is a meeting. The meeting allocates quantities according to the rules above. If there is no meetings quantities are the same as in the previous period.

At this point in the period all incumbents and potential entrants have the information required to make the decisions needed for this period. When we refer to the information set of period t we will be referring to the information available at this point in the period. Simultaneously, profits are allocated and all decisions are made (this includes entry, exit, and investment decisions, as well as decisions on whether to call a meeting for the following period).

Information Sets.

It is convenient to divide the (payoff relevant) information available to each agent into public and private information. If we let J^p_t be the information that is available to all incumbents and potential entrants at the time quantity decisions were made (i.e. after a meeting if there was one), then

$$J^p_t = \{\omega_{t-\xi(t)}, \hat{\xi}(t), \nu(\hat{\xi}(t))\} \in \mathcal{J}^p,$$

where,

- $\xi(t)$ is the number of periods that have passed since the meeting (if there was a meeting at period t then $\xi(t)=0$). To ease the computational burden we assume imperfect recall; i.e. firms can recall information from at most $\overline{\xi}$ periods. Thus we set $\hat{\xi}(t) \equiv min\{\xi(t), \overline{\xi}\}$. So at period t firms only recall the precise date of the last meeting if that meeting was less than $\overline{\xi}$ periods before, otherwise they just realize it was more than $\overline{\xi}$ periods ago. Note however that $\omega_{t-\xi(t)}$ is in the information set in every period since it was always used in period $t-1$.

- $\nu(\hat{\xi}(t)) \equiv (\nu_{t-\hat{\xi}(t)+1}, \ldots, \nu_t)$, with the understanding that if a meeting was called in the current period $\nu(\hat{\xi})$ is empty. As above firms recall at most the common shocks over the last $\overline{\xi}$ periods.

- $\mathcal{J}^p = \Omega^N \times \{1, \ldots, \overline{\xi}\} \times \{0,1\}^{\overline{\xi}}$, where $\omega \in \Omega$ and N is the maximum number of firms ever active. We invoke regularity conditions similar to those in Ericson and Pakes (1995), to insure that both N and $\#\Omega$ are finite.

$J_{i,t}$ will be firm i's information set in t (i.e., the public information and the firm's own state). If the firm is an incumbent then

$$J_{i,t} = \{\omega_{i,t}, J_t^p\} \in \Omega \times \mathcal{J}^p \equiv \mathcal{J},$$

while if it is a potential entrant $J_{e(t)} = (J_t^p, x^e)$, where x^e is its entry cost.

Strategies and Their Costs.

Recall that the following decisions are made simultaneously: exit or $\chi_{i,t} : (J_{i,t}) \rightarrow \{0,1\}$; entry or $\chi_t^e : (J_t, x^e) \rightarrow \{0,1\}$; calling a meeting, or $cm_{i,t} : (J_{i,t}) \rightarrow \{0,1\}$; and investment or $x_{i,t} : (J_{i,t}) :\rightarrow R^+$.

We have already specified the costs of entry and exit. The costs of investment, $c(x) = x$ and the cost of calling a meeting whenever there is more than one firms are given by

$$c(cm_i, cm_{-i}, \chi^e) = cm_i(FK + \Delta FK) + \{\sum_{j \neq i} cm_j > 0 \text{ or } \chi^e = 1\}[1 - cm_i]FK$$

where $\{\cdot\}$ is the indicator function which takes the value one if the "$\cdot$" condition is satisfied. We have assumed, as is done in our computations, that the cost of a meeting which reallocates quantities after an exit is zero. Of course if there is only one firm in the industry it can change its output without incurring the costs of calling a meeting.

3.3 Equilibrium.

Let $\mathcal{S} = \Omega^N \times \mathcal{J}^p$ and let s be its generic element. Note that an $s \in \mathcal{S}$ defines a tuple $(J^p, J_1, \ldots, J_{n(J)})$, where $J_i = J^p \times \omega_i$ provides the information available to the different incumbents, and (J^p, x^e) provides the information available to the potential entrant. We will say that J_i (or J^p) is a component of s if there is a (J^p, J_{-i}) such that $(J^p, J_i, J_{-i}) = s$. Adopting our MPE definition in Section 2.1 we define an MPE for our example as follows.

Definition: Markov Perfect Equilibrium.

An equilibrium is a tuple:

- $W = \{(W(\eta|J_i, cm),\ \forall J_i \in \mathcal{J},\ \eta \in \{0,1\},\ cm \in \{0,1\}\}$ and a $V^{en}(J^p)$, $\forall J \in \mathcal{J}$ (below $V^{en}(\cdot)$ will represent the value of entry),
- Strategies, $(x(J_i), \chi(J_i), cm(J_i), \chi^e(J^p, x^e))$, for each J^p and J_i which is a component of an $s \in \mathcal{S}$,

that satisfy the following conditions.

$C1$. Strategies are optimal given $W(\cdot|J_i, cm)$ and $(V^{en}(J^p), x^e)$.

$C2$. $W(\eta|J_i, cm)$ and $V^{en}(J^p)$ are the expected discounted value of net cash flows given the equilibrium strategies.

That is

$$W(\eta|J_i, cm) = \beta E\left[-c(cm, cm_{-i}, \chi^e) + V(J_i') \mid \eta, J_i\right],$$

where

$$V(J_i') = \pi(J_i') +$$
$$max\left\{\beta\phi, max_{cm} sup_x \left[W(1|J_i', cm)p(x, \omega_i') + W(0|J_i', cm)(1 - p(x, \omega_i')) - x\right]\right\},$$

and

$$V^{en}(J) = E\left[-FK + \beta V(J'_{\omega(en)})|J^p\right],$$

where $\omega(en)$ is the ω at which a new entrant begins operations, and the perceived probability distributions used to take expectations are consistent with equilibrium behavior. ■

Note that the expectations in this definition are taken using the density $p(J_i'|J_i, \eta, cm)$, for all $J_i \in \mathcal{J}$, $\eta \in \{0, 1\}$ and $cm \in \{0, 1\}$ which provide the perceived probabilities of J_i' given any (J_i, η_i, cm). Another way to state the equilibrium conditions is to state that these probabilities are consistent with the equilibrium strategies of all competitors.

Issues of existence and uniqueness for complete information models within the Ericson and Pakes (1995) framework are discussed in Doraszelski and Satterthwaite (2003). Here we simply assume existence and consider issues associated with analyzing the equilibrium. As noted $\mathcal{S}$ has a finite number of elements. Consequently any equilibrium generates a finite state Markov chain on $\mathcal{S}$. In this context there are a number of implications of our definition that will allow us to simplify the problem of calculating equilibrium policies.

- First note that were we to know the $\{W(\eta|J_i, cm(J_i))\}$ for each J_i component of $s \in S \subset \mathcal{S}$ we would have enough information to compute equilibrium policies from any point in S.

- Second given equilibrium play, the objective (or empirical) distribution of outcomes, say $p^{e,i}(J_i'|J_i, \eta_i)$, are random draws from the equilibrium $p(J_i'|J_i, \eta, cm(J_i))$. This implies that the empirical distribution from points that are visited infinitely often converges to $p(J_i'|J_i, \eta, cm(J_i))$, the equilibrium distribution of outcomes.

- Finally, every sample path will, in finite time, wander into a recurrent subset of points, say $\mathcal{R} \subset \mathcal{S}$, and once in $\mathcal{R}$ will stay in it forever.

Together the first and third points imply that to analyze subgames from a point in $\mathcal{R}$ we only need to know the $\{W(\eta|J_i, cm(J_i))\}$ for each J_i component of $s \in \mathcal{R}$. Further the second point enables us to test whether a candidate sequence of $\{W(\eta|J_i, cm(J_i))\}$ on a subset of the space obtained from the algorithm are in fact equilibrium $\{W(\eta|J_i, cm(J_i))\}$. We simply substitute the empirical distribution of outcomes for the $p(J_i'|J_i, \eta, cm)$ on the right hand side of conditions $C1$ and $C2$ above and test for the equality signs in those conditions[24].

---

[24] Strictly speaking this can only be done for points that are "interior" points of the recurrent class, where, as in Pakes and McGuire, 2001, interior points are points at which

We now provide a simple reinforcement learning algorithm that can be used to approximate the sufficient statistics needed for equilibrium play. We should note, however, that an alternative approach would have been to view the algorithm itself as the way players learn the statistics needed to choose their policies, and justify the output of the algorithm in that way. A reader who subscribes to the latter approach may be less interested in the section of the paper on testing (which follows the computational section).

4 Computing the Equilibrium

We begin with an overview of how our computational algorithm works. We then provide the details needed to construct the algorithm.

4.1 Overview of the Algorithm

Consider the Bellman equation with the information available just before all decisions are made and profits allocated. This will define equilibrium values and policies.

If we let $\pi(J_{i,t}) \equiv \pi(\omega_{i,t}, q(\omega(t-\xi_t)))$, and omit the (i,t) index for notational convenience, then the Bellman equation

$$V(J) = \pi(J) + max\left\{\beta\phi; max_{cm\in\{0,1\}} sup_{x\geq 0}\left[-x + \sum_{\eta} W(\eta|J,cm)p(\eta|x,\omega)\right]\right\} \tag{6}$$

where

$$W(\eta^*|J,cm) \equiv E_{(J')}\left\{\beta V(J') - c(cm, cm_{-i}, \chi^e)|\eta' = \eta^*, J, cm, \chi = 1\right\}. \tag{7}$$

Writing the Bellman equation in this way makes it easy to see that were we (or the firm's) to know the values for $\{W(\eta|J,cm)\}$, they would be sufficient for calculating optimal policies and the values associated with them. That is the $\{W(\eta|J,cm)\}$ are sufficient statistics for the decision problem. Consequently our algorithm will look for an efficient way of computing them.

---

agents cannot communicate to points outside of the recurrent class no matter which among the feasible strategies are played (by definition they can only transit to points inside the recurrent class if they play equilibrium strategies). Under our assumptions the data itself should enable us to separate the recurrent points into interior and boundary points.

Note that were we to compute the fixed point defining equilibrium behavior iteratively, i.e. if we were to compute $V(\cdot)$ at each iteration of a successive approximation routine, we would have to evaluate

$$E_{(J')}\left\{V(J')|\eta' = \eta^*, J, cm, \chi = 1\right\}$$

explicitly at each point at each iteration. This would require explicit calculation of *posterior probabilities* of the form

$$Pr(\omega_{-i,t} = \omega_{-i}^*|J_{i,t}, cm_{i,t}, \chi_{i,t} = 1)$$

at each point. These probabilities would have to be calculated recursively and kept in memory. Moreover the cardinality of the set of $J_{i,t}$ (and hence the number of distributions that would have to be kept in memory) is subject to the curse of dimensionality (it would increase exponentially in ξ, the number of active firms, the cardinality of Ω, etc.).

We present a computational algorithm which never requires us to calculate this expectation, and hence never requires us to calculate and retain posterior probabilities. Instead we will compute the $W(\eta|J, cm)$ in the equation above iteratively using techniques analogous to those used in the stochastic approximation (or reinforcement learning) literature (see the references above).

The stochastic approximation algorithm has an initiation procedure which provides initial values of $\{W(\eta|J, cm)\}$ for all η and $cm \in \{0, 1\}$ for different locations (say L) of the algorithm, where a location designates the public and the private information sets of all firms active. Starting at any L^0 the algorithm calculates policies for all agents by assuming that the $\{W^0(\eta|J, cm)\}$ are the true continuation values or $\{W(\eta|J, cm)\}$. Note that we do not need to compute an integral over possible future values in order to obtain policies. This is the second computational advantage of the algorithm (and the computational burden of the integral is also subject to the curse of dimensionality).

Given these policies, and computer generated draws for the realizations of all random variables which determine outcomes conditional on the policies, the algorithm moves to a new location, say L^1. The draws are then treated as if they are random draws from the true distribution of outcomes given the policy. Since $W(\eta|J^0, cm)$ can be constructed as an expectation over these distributions, the draws themselves are used to update our estimate of them, i.e. to obtain $W^1(\eta|J^0, cm)$. The process is then continued iteratively from L^1.

Since by updating our estimates of $W(\cdot|J^0, cm)$ we are updating a mean, if the policies converge so should our estimates of the $\{W(\cdot)\}$ (and vice versa). Moreover since the estimate of the mean is a sample average we expect increasingly accurate estimates of the $W(\cdot)$ from a particular location the more times we visit that location. The location's visited repeatedly will eventually converge to a recurrent class of locations of the Markov Process generating $W(\cdot)$. This is the third major computational advantage of the algorithm. Not only does it do away with the problem of computing posteriors and using them to form expectations, it also need not obtain accurate estimates of the $W(\eta|J, cm)$ at locations which are not in $\mathcal{R}$.

We now provide a formalization of the algorithm and then come back to the issue of testing whether a candidate set of $W(\cdot)$ outputted by our algorithm do in fact satisfy the equilibrium conditions.

4.2 Notation for the Stochastic Algorithm.

The iterative stochastic algorithm makes repeated use of the relationship between continuation values and the $\{W(\eta|J, cm)\}$. Consequently it will be helpful if we begin with those relationships. Rewrite the Bellman equation as

$$V(J_{i,t}) = max\left\{\beta\phi + \pi(\omega_{i,t}, q(\omega(t-\xi_t)); max_{cm_{i,t}\in\{0,1\}}V^c(J_{i,t}, cm_{i,t})\right\}, \quad (8)$$

where

$$V^c(J_{i,t}, cm_{i,t}=1) = \pi(\omega_{i,t}, q(\omega(t-\xi_t))) \quad (9)$$
$$+max_{x\in R^+}\{E[-x - c(cm_i=1, cm_{-i}, \chi^e) + \beta V(J_{i,t+1})]|J_{i,t}, x, cm_{i,t}=1\},$$

and

$$V^c(J_{i,t}, cm_{i,t}=0) = \pi(\omega_{i,t}, q(\omega(t-\xi_t)) \quad (10)$$
$$+max_{x\in R^+}\left\{E\left[-x - c(cm_i=0, cm_{-i}, \chi^e) + \beta V(J_{i,t+1})|J_{i,t}, x, cm_{i,t}=0\right]\right\}.$$

Note that if

$$W(\eta|J_{i,t}, cm_{i,t}=1) \equiv$$

$$E[(\beta V(J_{i,t+1}) - c(cm_i = 1, cm_{-i}, \chi^e)|\eta_{i,t+1} = \eta, J_{i,t}, cm_{i,t} = 1]$$

for $\eta \in \{0, 1\}$, then

$$V^c(J_{i,t}, cm_{i,t} = 1) = \pi(\omega_{i,t}, q(\omega(t - \xi_t))) \tag{11}$$

$$+max_{x \in R^+}[-x + W(1|J_{i,t}, cm_{i,t} = 1)p(x, \omega_i) + W(0|J_{i,t}, cm_{i,t} = 1)(1 - p(x, \omega_i))].$$

Similarly if

$$W(\eta|J_{i,t}, cm_{i,t} = 0) \equiv$$

$$E\left[(-c(cm_i = 0, cm_{-i}, \chi^e) + \beta V(J_{i,t+1}))|\eta_{i,t+1} = \eta, J_{i,t}, m_{i,t} = 0\right]$$

for $\eta \in \{0, 1\}$, then

$$V^c(J_{i,t}, cm_{i,t} = 0) = \pi(\omega_{i,t}, q(\omega(t - \xi_t))) \tag{12}$$

$$+max_{x \in R^+}[-x + W(1|J_{i,t}, cm_{i,t} = 0)p(x, \omega_i) + W(0|J_{i,t}, cm_{i,t} = 0)(1 - p(x, \omega_i))].$$

Finally the Bellman equation for a potential entrant is

$$V_e(J_t) = \beta E[V(\omega^e - \nu_{t+1}, J_{t+1}) - FK\{\# \, of firms > 0\}|\chi^e_t = 1, J_t].$$

4.3 Details of the Stochastic Algorithm.

An iteration, which will be indexed by k, is defined by a location, say L_k, and by a memory, which will be designated M_k.

4.3.1 Storage and Policies at Iteration k

The location is defined as a tuple

$$L^k = \{J^k, \omega^k_1, \ldots, \omega^k_{n(J(k))}\}$$

where ω^k_j is the current ω of the j^{th} largest firm in the $\omega_{t-\xi(t)}$ specified in J^k, and as before,

$$J^k = \{\omega^k_{t-\xi(t)}, \hat{\xi}^k_t, \nu^k(\hat{\xi}(t))\}.$$

There is the possibility of storage in memory at each possible L. Distinct objects are stored at J. Further for each J there can be items stored at each of the triples, $\{J, j, \omega\}$ for $j = 1, \ldots, n(J)$ and $\omega \in \{1, \ldots, \Omega\}$ (of course at any iteration some of these will have no information stored). I.e. for each J^k we begin with the largest firm in the ω tuple defining J^k, find out its current ω and list objects under the triple $(J^k, 1, \omega)$, then continue and store different objects under the triples $(J^k, 2, \omega)$ and so on.

The items stored are as follows.

- For each J^k we will have $M(J^k)$ stored at J^k where $M(J^k)$ contains
  - The number of times we have visited J^k, or $h^k(J^k)$, and if $h^k(J^k) > 0$
  - $V_e^k(J)$ the k^{th} iteration's estimate of the value of entry at J
  - $q(J^k, j)$ for $j = 1, \ldots, n(J^k)$.

  Note that if $h^k(J^k) = 0$, nothing is in memory for that point.

- For each $\{J^k, j, \omega\}$, we have $M(J^k, j, \omega)$, or the information stored in memory at (J^k, j, ω) as
  - $h^k(J^k, j, \omega)$ the number of times we have hit (J^k, j, ω), and if $h^k(J^k, j, \omega) > 0$
  - $W^k(\eta | J^k, j, \omega, cm)$ for $cm \in \{0, 1\}$ and $\eta \in \{0, 1\}$.
  - $\pi(J^k, j, \omega)$[25].

4.3.2 Updating and Initialization.

We require the following updates.

- Update $L^k = \{J^k, \omega_1^k, \ldots, \omega_{n(J(k))}^k\} \rightarrow L^{k+1} = \{J^{k+1}, \omega_1^{k+1}, \ldots, \omega_{n(J(k+1))}^{k+1}\}$
- Update $M(J^k)$. Here we update only $V_e^k(J^k)$ and $h^k(J^k)$, see below.

---

[25] It might also be efficient to store $V^k(J^k, j, \omega)$, and $x^k(J^k, j, w)$, $\chi^k(J^k, j, w)$, $m^k(J^k, j, w)$, rather than compute them, as needed from the equations above.

- Update $M(J^k, j, \omega)$, for $j = (1, \dots, n(J^k))$. Here we update

  $$W^k(\eta|J^k, j, \omega, cm) \rightarrow W^{k+1}(\eta|J^k, j, \omega, cm)$$

  for $\eta \in \{0, 1\}$ and $cm \in \{0, 1\}$, and $h^k(J^k, j, \omega)$, see below.

In doing these updates we will use the operator $V(\cdot|W)$ defined as

$$V(J, j, \omega|W) \equiv max\left\{\beta\phi + \pi(J, j, \omega), V^c(J, j, \omega|W)\right\},$$

where

$$V^c(J, j, \omega|W) \equiv$$

$$\pi(J, j, \omega) + max_{cm\in\{0,1\}} max_x[-x + W(1|J, j, \omega, cm)p(x, \omega) + W(0|J, j, \omega, cm)(1 - p(x, \omega))].$$

While we do the update we initialize if required. That is

- If $h^k(J^{k+1}) = 0$ compute $q(\omega^{k+1}) = s^N(\omega)Q^c(\omega)$ (from equations 2 and 4) and put it in memory for J^{k+1} . Also if we have to initialize we set

  $$V_e^k(J^k) = W^k(0|J^k + I(\omega_e), j(\omega_e, J^k + I(\omega_e))), \omega_e, cm = 1),$$

  where here and below

  $$\omega^{k+1} + I(z)$$

  adds an $\omega = z$ to the ω^{k+1} vector, and then reorders it in the natural order, and where $j(z, J)$ provides the order of the $\omega = z$ element in the ω vector defined by J. If $W^k(0|J^k + I(\omega_e), j(\omega_e, J^k + I(\omega_e))), \omega_e, cm = 1)$ initialize it as specified below.

- If $h^k(J^{k+1}, j, \omega) = 0$, calculate $\pi(J, j, \omega) = \pi(\omega, s^N(\omega)Q^c(\omega))$ (from equation 3) and put in memory. Also initialize

  $$W^1(\eta|J, j, \omega, cm = 0) = \pi(J, j, \omega + \eta)/([1 - \beta], \text{ for } \eta = \{0, 1\},$$

  and

  $$W^1(\eta|J, j, \omega, cm = 1) =$$

  $$\beta\pi(J - I(j, J) + I(\omega + \eta), j(J - I(j, J) + I(\omega + \eta)), \omega + \eta)/([1 - \beta],$$

  for $\eta = \{0, 1\}$.

- After initializing W-s, compute optimal policies given the W-s.

4.3.3 Update 1: Policies.

Get the realization of $\widetilde{x}^e$.

- Determine whether $\chi_e^k(J^k) = 1 \Leftrightarrow V_e^k(J^k) \geq x_e$.
- Choose $x(J^k, j, \omega, cm)$ as

$$argmax_x[-x + \sum_\eta W^k(\eta|J^k, j, \omega, cm)p(\eta|x, \omega)],\ for\ cm \in \{0,1\},$$

- Calculate

$$V^c(J^k, j, \omega, cm) =$$

$$\pi(J^k, j, \omega) - x^k(J^k, j, \omega, cm) + \sum_\eta W^k(\eta|J^k, j, \omega, cm)p(\eta|x(J^k, j, \omega, cm), \omega)$$

  for $m \in \{0,1\}$.

- Calculate $m(J^k, j, \omega) = argmax_{cm\in\{0,1\}} V^c(J^k, j, \omega, cm)$ and $V^c(J^k, j, \omega) = max_{cm\in\{0,1\}} V^c(J^k, j, \omega, cm)$
- Calculate $\chi(J^k, j, \omega) = 0 \Leftrightarrow V^c(J^k, j, \omega) \leq \beta\phi + \pi(J^k, j, \omega)$

Now we have all the policies. For each (J^k, j, ω) we have calculated $x(J^k, j, \omega, cm)$ and then $V^c(J^k, j, \omega, cm)$ for $cm \in \{0,1\}$; then $cm(J^k, j, \omega)$, and finally $\chi(J^k, j, \omega)$.

4.3.4 Update 2: Finding The New Location.

To obtain the new location we need the draws from distributions determined by the policies just calculated. So we make the draws first, keep them in the working file, and use them below.

- Draw ν^{k+1}. Here $\nu^{k+1} = 1$ with probability δ and $\nu^{k+1} = 0$ with probability $1 - \delta$.

- For each (J^k, j, ω) such that $\chi(J^k, j, \omega) = 1$, use $x(J^k, j, \omega, cm(J^k, j, \omega))$ and ω to draw η_j^{k+1} and calculate $\omega^{k+1}(J, j) = \omega^k(J, j) + \eta_j^{k+1} - \nu^{k+1}$ (note $\chi^k(J_j^k) = 0 \Rightarrow \omega_j^{k+1} = 0$). Note $\eta_j^{k+1} = 1$ with probability $A(\omega_j)x(cm)_j/(1 + A(\omega_j)x(cm)_j)$ and $\eta_j^{k+1} = 0$ with probability $1/(1 + A(\omega_j)x(cm)_j)$ where $A(\omega_j)$ is decreasing in ω_j.

Now we have all the needed random draws. This enables us to update L^k.

- If $\sum_i [1 - \chi(J_i^k)] = 0$ and $\sum_j cm(J_j^k) = 0$ and $\chi_e(J^k) = 0$, then

$$J^{k+1} = \{\omega_{k-\xi(k)}, \hat{\xi}(k+1) = min[\hat{\xi}(k) + 1, \overline{\xi}], \nu(\hat{\xi}(k+1))\}$$

- Otherwise form ω^{k+1} by taking $\omega^{k+1}(J^k, j)$ for each j at which $\chi(J^k, j) = 1$ and $\omega^e - \nu^{k+1}$ if $\chi^e(J^k) = 1$, and ordering the result (in the natural order)

$$J^{k+1} = \{\omega^{k+1}, 0, 0\},$$

- $L^{k+1} = (J^{k+1}, \omega^{k+1})$.

4.3.5 Update 3: $M(J^k)$.

First set

$$h^{k+1}(J^k) = h^k(J^k) + 1.$$

Next set

$$V_e^{k+1}(J^k) - V_e^k(J^k) = [h^k(J^k) + 10]^{-1}$$

$$[\beta(V(J^e(\omega^{k+1}) - Z(J_e, J_e', cm = 0)), j(\omega_e - \nu^{k+1}, J^e(\omega^{k+1})), \omega_e - \nu^{k+1}) | W^k) - V_e^k(J^k)]$$

where

$$J^e(\omega^{k+1}) = J^{k+1}$$

if $\chi_e^k = 1$ and

$$J^e(\omega^{k+1}) = (\omega^{k+1} + I(\omega_e - \nu^{k+1}), 0, 0)$$

otherwise.

4.3.6 Update 4: $M(J^k, j, \omega)$.

First we update
$$h^{k+1}(J_j^k) = h^k(J_j^k) + 1.$$

Next we have to update $W^k(\eta|J^k, j, \omega, cm)$ for $\eta \in \{0, 1\}$ and $cm \in \{0, 1\}$. The update can differ with

$$\chi_j^k, cm_j^k, \eta_j^{k+1}, \{\sum_{i \neq j} cm_i^k > 0\}, \{\sum_{i \neq j} \chi_i^k > 0\}, \text{ and } \chi_e^k.$$

Thus for each j there are four updates and the way that update is made could differ for each of 2^6 possible outcomes. We show below however that they reduce to four cases, three of which can be calculated from a single formula.

All updates will be of the form

$$W^{k+1}(\eta|J^k, j, \omega, cm) - W^k(\eta|J^k, j, \omega, cm) = [h^k(J_j^k) + 2]^{-1}[V^\dagger(*) - W^k(\eta|J^k, j, \omega, cm)],$$

and what we have to do is provide the form of $V^\dagger(*)$.

Case 1: We evaluate a situation in which there is a meeting, i.e. either $cm = 1$, or $\sum_{i \neq j} cm_i^k \geq 1$, or $\chi_e^k = 1$, or $\sum_{i \neq j}[1 - \chi_j^k] \neq 0$.

There are two possible J^* in this case, one for $\chi_j^k = 1$ and one for $\chi_j^k = 0$, and the $V^\dagger(*)$ will depend on J^*. If $\chi_j^k = 1$, then

$$J^*(\cdot) = ((\omega^{k+1} - I(\omega_j^k + \eta_j^{k+1} - \nu^{k+1}) + I(\omega_j^k + \eta - \nu^{k+1}), 0, 0).$$

If $\chi_j^k = 0$, then

$$J^*(\cdot) = ((\omega^{k+1} + I(\omega_j^k + \eta - \nu^{k+1})), 0, 0).$$

In either case

$$V^\dagger(*) = \beta V(J^*(\cdot), j(\omega_j^k + \eta - \nu^{k+1}, J^*), \omega_j^k + \eta - \nu^{k+1}|W^k) - c(cm, cm_{-i}, \chi^e),$$

where the three subcases correspond to the three different possible values of $c(\cdot)$. Recall that these are: $cm = 1$ in which case $c(\cdot) = FK + \Delta FK$; $cm = 0$ and either $\chi_e^k = 1$ or $\sum_{i \neq j} cm_i > 0$, in which case $c(\cdot) = FK$; and $cm = \chi_e^k = 0$ and $\sum_{i \neq j} cm_i = 0$ in which case $c(\cdot) = 0$.

Case 2: In the position we are evaluating there is no meeting at all, i.e. $cm = 0, \chi_e^k = 0, \sum_{i \neq j} cm_i^k = 0$ **and** $\sum_{i \neq j}[1 - \chi_j^k] = 0$**.**

Note that these are the cases when the new states are not fully revealed in the state we are evaluating. Here it does not matter whether χ_j^k is. Either way

$$V(*) = \beta V[(\omega_{k-\xi(k)}, min(\hat{\xi}(k) + 1, \bar{\xi}), \nu(\hat{\xi}(k) + 1), j, \omega_j^k + \eta - \nu_j^{k+1}|W^k]. \ \blacksquare$$

Note that After Updating the $W^k(\cdot)$ then we can update $V^k(\cdot)$ where

$$V^{k+1}(\cdot) = V(\cdot|W^{k+1})$$

using the $V(\cdot)$ operator defined above, as needed.

4.4 Computational Comparison.

Standard iterative computational algorithms for equilibrium fixed points keep values and policies in memory for each point in the state space, update these values sequentially at every iteration, and stop the algorithm when the values and policies in two successive iterations are equal. Their computational burden is determined as the product of; (i) the number of points in the state space, (ii) the burden of updating each point at each iteration, and (iii)the number of iterations (for more detail, and an extensive discussion of procedures for simplifying the computational burden, see Judd, 1998). As noted in Pakes and McGuire (2001), the use of stochastic approximation algorithms reduces the burden of computing Markov Perfect equilibrium in dynamic games in two ways: the algorithm eventually focuses in on a recurrent class of points thus reducing the number of points updated, and the computational burden at each point evaluated changes from the burden involved in computing integrals over future values to the burden involved in updating averages of past values.

When we allow for asymmetric information the computational burden updating each point in an iterative algorithm increases *dramatically.* The expectation over the future required for the updated continuation value is taken with a probability distribution which itself must be computed as a ξ-fold convolution of more primitive distributions, and each of these distributions depends on the policies relevant at *alternative* points in the state space. The fact that those probabilities depend on policies at other states will require us to either increase the memory for each state substantially, or to search and retrieve information from different states each time we update for a particular state. Given these policies the computational complexity of the ξ-fold convolution increases exponentially in ξ. In rather stark contrast, the updating burden of the stochastic algorithm remains the same; it still only need to update averages.

5 Numerical Results.

We proceed as follows. First we work with our base case. We provide the statistics which tests the MPE conditions for the run that computed the equilibrium for that case and describe the equilibrium these parameter values generate. Next we engage in a "comparative dynamic" analysis of the impacts of changes in FK. Lastly we compare these results to those obtained from running our model with two different institutional settings; (i) a full information case in which firms observe the cost position of their rivals, and (ii) a model in which side payments are feasible and production is done only by the most efficient firm.

5.1 The Test Statistics.

Recall from section 2.2 that the test uses the policies outputted by the algorithm to simulate sample paths, and then compares the average value of the simulated paths at each point to the values outputted by the algorithm (after accounting for sampling error in the simulated paths). The test statistic itself has the interpretation of an $\mathcal{L}^2(P)$ norm in the percentage deviation between the simulated and estimated values, where the weights in P are determined by the frequency with which the different points are visited (as an approximation to their probability in the invariant distribution on the recurrent class of points).

As noted by varying the number of periods simulated or the stopping rule (the τ in section 2.2) we can construct many different tests, and the test with $\tau \equiv 1$ is the stochastic analogue of the stopping rule typically used in iterative algorithms. Figure 1 graphs the values of both the $\tau = 1$ test, and a second test which does not rely on simulation but is not generally available for dynamic games with asymmetric information. In our example there are a subset of the points in $\mathcal{R}$, the points at which there is a meeting (or $\xi = 0$), in which there is no asymmetric information. At these points we can calculate the discounted future values that the algorithm's policies imply without computing posterior distributions, and hence we can mimic the exact test (i.e. a test that does not require simulation) used in models without asymmetric information (see Pakes and McGuire, 2001, and the literature they cite). The solid line provides the $\mathcal{L}^2(P)$ norm of the percentage differences of the exact test at the points with full information[26].

Each test statistic was computed at ten million iteration intervals, starting after the first ten million, and continuing until two billion. The solid line in the figure shows the "exact" test statistic. It starts at .1% after ten million iterations, and declines rather rapidly between the ten and thirty fifth million iterations where it takes on a value of .01%. The statistic does decline further as we increase the number of iterations, but the rate of decline slows continually and becomes almost imperceptible. It takes until seven hundred annd fifty million until it reaches .001%, and at two billion it is .0007%.

This test does not provide information on the closeness of the estimated values to the values implied by our policies at points at which there is no meeting. The second test, which is exactly the test described in section 2.2 (with $\tau = 1$), uses simulation to construct comparisons for all points with weights given by the points' frequencies in the last ten million draws. These results are more initially more "jumpy" (as one would expect from a stochastic test), and start much higher, at 1.9%. The statistic falls to .1% at three hundred million. However it falls at a much slower pace after that and takes until one billion eight hundred million to fall to .01% [27].

It seems quite clear that we are converging to a fixed point and the only question is at what level of accuracy to stop the iterations. We have done many preliminary runs with different specifications and done an assortment

---

[26] In fact we computed several other tests, but we will suffice here with a description of how these two test statistics behaved as the results on the other tests were similar.

[27] The fact that the norm is larger for the second test could result from a number of different phenomena. There could be; a poorer fit at the points where there is assymetric information, some inaccuracy generated by the random number generator, or noise in our measure of sampling variances. Two other points are worth noting here. First these are uncenterred differences. The actual correlation between the two discounted values was noticeably higher, with values exceeding .99999. Second we did not worry about boundary points in the tests for two reasons. First they were hit very infrequently, and second we initiate all points at very high values. As a result the points connected to the points on the boundary tend to be assigned higher values than their equilibrium values.

of auxiliary tests (e.g. comparing outputs from different runs, looking at subsequent changes in policies, etc.). As long as we stopped the algorithm when the second test was under .1%, all the tests indicated there was no reason to increase the number of iterations further.

5.2 The Benchmark Case

Summary statistics for several runs are presented in Table 1 and the ergodic distribution is given in Table 2. When $FK = 20$, which will be our benchmark case, the cost of renegotiation is about 57% of average net profits per firm per period.

With $FK = 20$ the industry is almost always a duopoly (see the second column in Table 1) with the same two firms operating for long intervals (on average for fifty-four periods), but with frequent renegotiations of market shares. 84% of the time they renegotiate market shares within four periods of setting them (see Table 4). When the two firms have very low $\omega's$ the length of time before a reallocation gets smaller, and the reallocation is almost always triggered by entry (see Figure 2 and Table 3). The three firm equilibria that result are very unstable; over 80% of the time one of these three firms exits within three periods. If the two remaining firm both have successful research they typically begin a long sample path in which they are the only competitors. If one firm grows and the other does not, the situation also becomes unstable and the low ω firm often exits. A temporary monopoly can result, but it is very temporary, its average duration being 1.8 periods, and is invariably broken by a second entrant, who is often followed by a third. The process then restarts itself.

As noted there are long periods when the same two active firms share the market, so in that sense the market structure is stable. However this hides the intense investment competition between the two incumbents, a competition that results in fairly frequent demands to renegotiate quantity allocations. At high $\omega's$ the value function becomes concave, and since the gains from investment are related to the slope of that function, high ω's generate less investment, and as a consequence less variance in the outcomes of the investment competition. However, as can be seen from Figure 2, even when both firms have very high ω's, the average interval of time between renegotiations is under four periods.

Meetings are typically called after at least one of the firms has improved its ω. As a result the public ω (which recall is the value of ω at the time of a meeting) is, on average, larger than the current ω (which is typically not observed by the participants, see table 1). Interestingly, in over 40% of the duopoly renegotiation periods the firm calling the meeting does not gain from it (that is the asymmetric information often causes "mistakes"; see table 5). Of course the average gain of the firm that renegotiates is positive. Indeed on average the sum of the net profits of the two firms increases by between 2 and 3 percent when there is renegotiation (recall that the renegotiation facilitates a more efficient allocation of output).

5.2.1 Comparative Statics; Different Levels of FK

Recall that FK is the cost of a meeting. We examine the implications of varying it as one can think of this cost being related to the resources the antitrust authority allocates to monitoring the industry.

Perhaps the most striking finding is that consumer surplus increases with FK (see the last panel of Table 1). This is a result of price falling as FK increases, a phenomena which occurs despite the fact that increases in FK increase the cost of doing business, and hence generates equilibria with less firms operating. It is clear that even in standard Cournot games with asymmetric costs price can fall when we move from a duopoly to a monopoly (at least provided it is the high cost firm that exits). In a model with cost reducing investments, however, the tendency for price to fall is likely to be accentuated. This because a reduction in the number of firms typically increases the quantities per firm which, in turn, increases investment incentives which causes further cost reductions.

When FK increases from 10 to 20 there is only a very small reduction in the distribution of the number of firms in equilibrium. Consequently the direct effect of increasing FK on the costs of doing business is the dominant effect of the change in FK on producer surplus and it falls. However when FK moves from 20 to 30, there is a large increase in the number of monopoly periods. In monopoly periods positive outcomes of the investment process can be utilized in a more efficient way. This is because the monopolist can apply the cost reduction to the entire market and it can change quantities costlessly.

Note that this implies that once dynamics are taken into account both producer and consumer surplus are highest when $FK = 30$.

5.3 Institutional Variations.

There are several novel features of our benchmark model; the costly renegotiation required to change prices and market shares, the asymmetric information or lack of knowledge of rivals' costs, and the lack of an ability to coordinate on an efficient allocation of production at meetings when shares are changed (efficiency would require $s_i = 1$ for $i = argmin_i mc_i$). The comparative static exercise above was meant to clarify the role of the costs of renegotiation. We now consider exercises designed to clarify the role of asymmetric information and then to clarify the implications of the inability to coordinate on an efficient allocation of production.

5.3.1 Full Information (FI) vs. Asymmetric Information (AI).

We now compare our results to those of a model in which the firms current ω's are perfectly observed by all firms, however changes in market shares and/or prices still requires costly renegotiation. Descriptive results for the FI model with FK=20 are provided in the fourth column in Table 1.

The most obvious change when there is full information is that there is much less renegotiation, entry, and exit. This reflects the fact that firms will no longer call a meeting, enter, or exit when post facto this would not be beneficial to them. Of course now when meetings are called they will invariably result in larger average price changes. All this generates higher producer surplus in the FI then in the AI model.

Interestingly, however, consumer surplus is higher in the AI model. This is despite the fact that investment per firm is slightly lower. The explanation lies in the change in when renegotiation occurs. In the AI model the current ω is typically lower than the public ω because a firm will only call for a renegotiation when they have substantially increased their ω. Since prices are set in renegotiation periods, prices will be set when costs of production are low. In contrast in the FI model a firm will call for renegotiation also when there is no increase in their own ω but their rivals do not keep up with the outside alternative. Thus prices will also be set when costs are higher in that model.

5.3.2 Collusive market with Efficient Production

In our benchmark model firms negotiate on market shares and each firm does its own production. This is not a cost efficient allocation. An efficient allocation would have $s_i = 1$ for $i = argmin\ mc_i$. For e.g. in our benchmark model (with $FK = 20$) the average unit cost is 1.21 while if we allocated production efficiently average costs would go down by 15% (to 1.03; see the second column of Table 1).

We now consider a variation in the institutional setting in which production is done only by the most efficient firm. We allow side payments and allocate profits according to their Cournot market shares, as this should both reduce possible enforcement problems (an issue we explicitly *do not* deal with), and provide some idea of how much firms stand to gain from a more efficient static allocation mechanism in a setting with asymmetric information and costs of adjusting market shares. Again we assume that decisions regarding which firm produces and how profits are allocated are done only during costly meetings in which cost positions become public information.

There is a large difference in investment and renegotiation incentives in the model with an efficient allocation of production and side payments. Then when the firm which is producing increases its ω and there is no meeting that firm enjoys the profits from its cost savings times the entire market output. This induces more investment per firm (see column 5 table 1). Alternatively if the costs of the low cost firm increases and there is no renegotiation, it bears the loss from the increase in cost times the entire market output. Consequently the low cost firm is now more likely to call for a meeting when its ω falls then when it rises. Both these differences lead to lower prices and increases in consumer surplus (see the last two columns of table 1).

Of course the more efficient allocation of production also causes an increase in produce surplus. We note, however, that in the parameterization we use in the above comparison the market structure generated by the two models is similar. This was not the case when we increased FK to 30. In that case the change to the optimal allocation induced more entry, less monopoly periods, lower investment per firm and hence higher costs, and subsequently lower producer surplus. Indeed the fall in producer surplus with optimal allocations and FK=30 was higher than the increase in consumer surplus when FK=30, so that total surplus was larger in the model without the efficient allocation.

It is clear that there is a potential for large gains from more efficient allocations of output. On the other hand it is equally clear that whether or not these gains are realized will depend on their dynamic or market structural implications; that is on the investment entry and exit responses to the change in the nature of the output allocation.

6 Concluding Remark

We have presented a simple algorithm for computing an intuitive notion of MPE for finite state dynamic games with asymmetric information. The algorithm is relatively efficient in that it does not require; storage and updating of posterior distributions, explicit integration over possible future states to determine continuation values, or storage and updating of information at all possible points in the state space. To illustrate our algorithm we computed the equilibrium of oligopolistic industries with collusive interactions. This showed that parameters determining information flows can effect market structure and through market structure producer and consumer surplus. More generally our hope is that the framework for analyzing dynamic games with asymmetric information presented here will enable more realistic analysis of interactions among economic agents.

References

Athey S and Bagwell, K (2001) "Optimal Collusion with Private Information" *The Rand Journal of Economics*, Vol 32, pp. 428-465.

Athey S and K. Bagwell, K and C. Sanchirico (2004) "Collusion and Price Rigidity", *Review of Economic Studies*, (forthcoming).

Bertsekas, D. and J. Tsitsiklis (1996) *Neoro-Dynamics Programming.* Belmont, MA: Athena Scientific Publications.

Besanko, D. and U. Doraszelski (2004) "Capacity Dynamics and Endogenous Asymmetries in Firm Size" *The Rand Journal of Economics*, 35(1), pp. 23-49.

Cheong K.S. and K. Judd, (2003) "Mergers and Dynamic Oligopoly," *Journal of Economic Dynamics and Control*, forthcoming.

Compte, O., F. Jenny and P. Rey (2002); “Capacity Constraints, Mergers and Collusion”, *European Economic Review*, 46(1), pp.1-29.

de Roos, N. (2000), "A Model of Collusion Timing" Working Paper, Yale University.

Ericson R. and A. Pakes, (1995) ”Markov-Perfect Industry Dynamics: A Framework for Empirical Work” *Review of Economic studies*, 62, 53-82.

Doraszelski, U. and M. Satterthwaite, (2003) "Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity," mimeo Hoover Institution.

Fershtman C. and A. Pakes (2000), "A Dynamic Oligopoly with Collusion and Price Wars" *The Rand Journal of Economics*, 31(2) pp. 207-236.

Freedman, D. (1983) *Markov Chains.* New York: Springer Verlag.

Gowrisankaran, G. (1999) "A dynamic model of endogenous horizon mergers" *The Rand Journal of Economics* 71, pp. 51-63.

Harrington, J. (1989), ”Collusion among Asymmetric Firms: The Case of Different Discount Factors” *International Journal of Industrial Organization*, 7, 289-307.

Harrington, J. (2002), "Cartel Pricing Dynamics in the Presence of an Antitrust Authority" Johns Hopkins University Working Paper.

Judd, K. (1998), *Numerical Methods in Economics*, MIT Press, Cambridge, Mass.

Levenstein M. (1997), ”Price Wars and the Stability of Collusion: A Study of the Pre-World War I Bromine Industry”, *Journal of Industrial Economics*, pp. 117-147.

Markovich S. (2000), "Snowball - The Evolution of Dynamic Oligopolies with Network Externalities," mimeo, Tel-Aviv University.

Maskin, E. and J. Tirole (1992), “Markov Perfect Equilibrium" *mimeo*, Harvard University.

Maskin, E. and J. Tirole (2001), "Markov Perfect Equilibrium: Observable Actions", *Journal of Economic Theory*, 100, pp. 191-219.

Pakes, A. (1994); “Dynamic Structural Models: Problems and Prospects”, Chapter 5 in C. Sims (ed.) *Advances in Econometrics Sixth World Congress*, Vol. II, pp 171-260, Cambridge University Press.

Pakes, A. and P. McGuire (1994), ”Computing Markov Perfect Nash Equilibrium: Numerical Implication of a Dynamic Differentiated Product Model”, *The Rand Journal of Economics*, 25, pp. 555-589.

Pakes, A. and P. McGuire (2001), ”Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the “Curse” of Dimensionality”, *Econometrica* pp.1261-81.

Posner, R.A. (1976), *Antitrust Law*, The University of Chicago Press, Chicago.

Robbins, H. and S. Monroe (1951); “A Stochastic Approximation Technique”, *Annals of Mathematics and Statistics*, 22, 400-407.

Scherer, F. (1980), ”*Industry Market Structure and Economic Performance*”, Chicago:Rand McNally.

Schmalensee, R. (1987), ”Competitive Advantage and Collusive Optima” *International Journal of Industrial Organization* 5, pp. 351-367.

Appendix

Parameter Values: For our numerical analysis, we adopt the following parameter values.

- Inverse demand $P = 6 - 4Q$.
- Marginal cost is $mc(\omega_i) \equiv \frac{.7\exp[-\omega_i/3]}{1+\exp[-\omega_i/3]}$.
- Transition probabilities $p(x_{i,t}, \omega_{i,t}) = \frac{A(\omega_{i,t})x_{i,t}}{1 + A(\omega_{i,t})x_{i,t}}$, where $A(\omega_{i,t}) = 0.33\exp[-\omega_i/3]$.

Our benchmark case has FK=20, though we examines also different values of FK. We always let ΔFK be 10% of FK. We set K=0, while the exit fee itself $\phi = 50$. The entrant pays the entry cost, x_e and enters the following period at $\omega = 1$ minus the exogenous shock. The entry cost are such that $x_e + FK \sim uniform[50,100]$. The discount rate, $\beta = .95$, and the probability of the outside alternative moving up, $\delta = .65$. These parameters imply that we never observe more than 5 firms active (i.e. N=5), and that only ω values between 1 and 25 are ever observed (or $\omega = 25$).

Figure 1: Test Statistics.

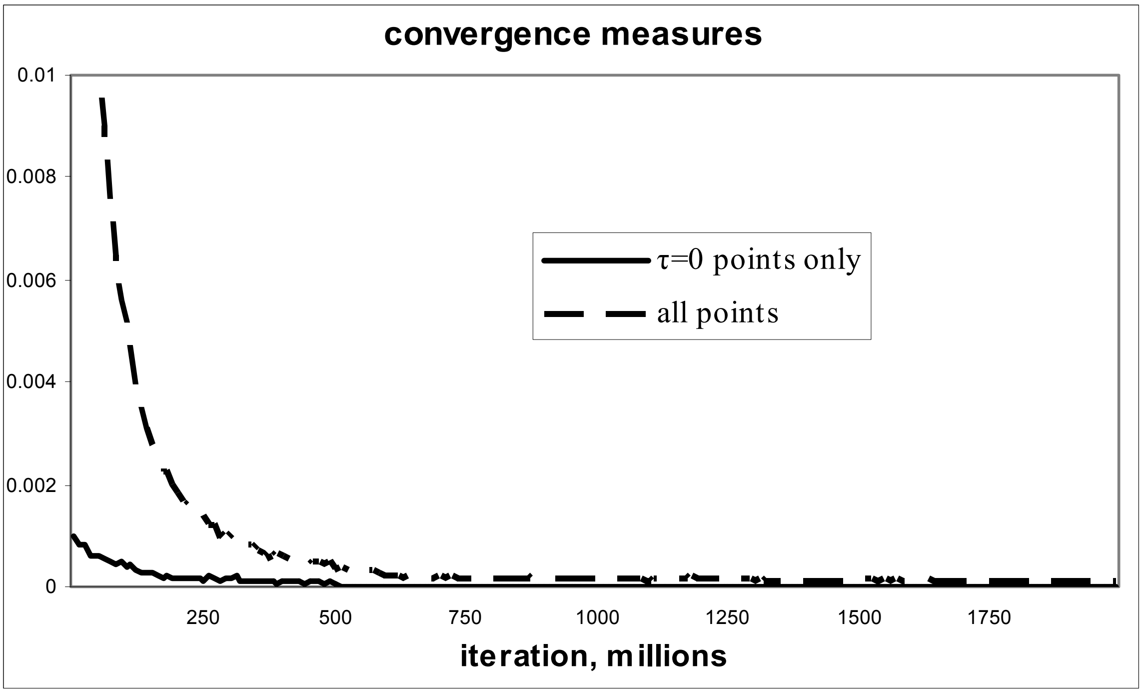
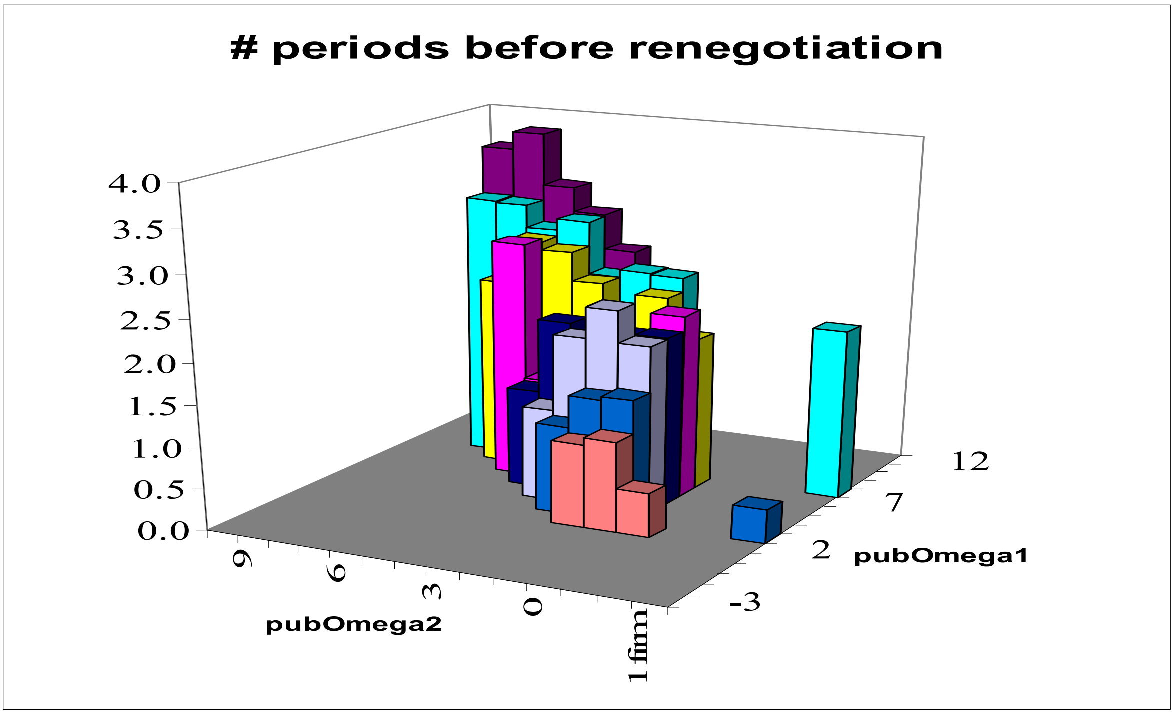


Table 1: Summary statistics: average value of characteristics. (Use last 400 millions iterations over three seeds).

	Asymmetric information			Full information	Optimal
	FK=10	FK=20	FK=30	FK=20	FK=20
Price	3.60	3.59	3.53	3.60	3.46
Q per firm	27.78	29.28	40.87	28.57	29.86
Total Q	60.07	60.26	61.64	60.05	63.38
Entry/exit/# firms					
% Entry	1.69	2.16	1.86	1.88	2.02
% Exit	1.57	1.99	1.76	1.81	1.84
% Entry&Exit	0.31	0.63	0.61	0.74	0.62
avg # firms	2.16	2.06	1.51	2.10	2.12
% 1 firm active	0.05	1.79	51.29	0.74	0.39
% 2 firms active	87.12	92.29	47.01	90.83	90.10
% 3 firms active	10.13	4.64	1.38	5.98	7.17
Renegotiation stats					
total % renegotiation	41.96	35.58	66.39	22.02	35.43
% Meeting (called at cost FK)	39.01	31.11	13.09	18.72	32.18
% meeting or entry	40.83	33.35	14.97	20.66	34.33
Omegas, investment, costs					
public omegas	4.69	4.77	5.11	4.69	4.86
current omegas	4.63	4.69	5.10	4.68	4.77
public-current	0.06	0.08	0.01	0.01	0.09
public omega of leader	5.62	5.66	5.71	5.59	5.93
current omega of leader	5.50	5.51	5.67	5.48	5.80
public omega of 2nd firm	4.14	4.06	4.01	4.03	4.13
current omega of 2nd firm	4.12	4.03	4.06	4.10	4.06
Investment per firm	27.79	28.26	33.49	28.27	29.54
Total investment	60.10	58.17	50.52	59.41	62.70
Avg public unit costs	1.19	1.18	1.07		
avg unit costs	1.22	1.21	1.07	1.21	0.96
efficient unit costs	1.03	1.03	0.98	1.03	0.96
Total cost of production	70.48	70.68	63.67	70.61	58.97
efficient total cost	59.55	59.93	58.50	59.90	58.45
Profits, consumer/producer surplus					
Gross profit per firm	66.68	70.20	101.33	68.61	75.11
avg net profit	34.61	34.96	61.74	36.35	38.49
Gross total profits	144.21	144.50	152.81	144.20	159.45
producer surplus	74.86	71.97	93.09	76.39	81.70
Consumer surplus	72.87	73.29	76.64	72.74	80.93

Figure 2: **Number of periods before renegotiation (baseline case).**



Note that the area which has zero # periods in the above figure is an area which is not in the ergodic distribution.

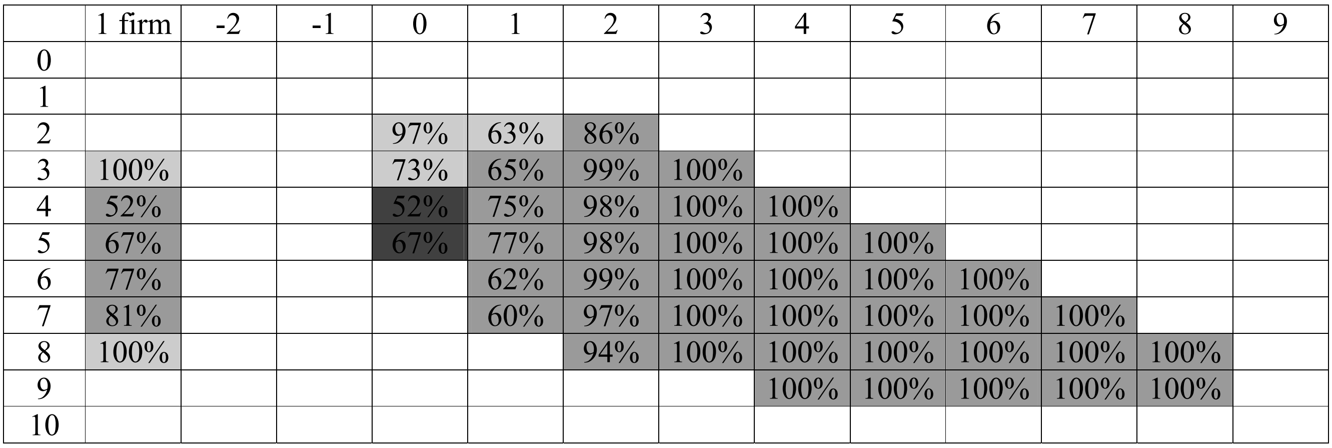
Table 2: **Ergodic distribution for the benchmark case: public omegas (last 100m iterations, 1-2 firms active, baseline parameterization, these points account for 93.7% of all periods).**

w1\w2	1 firm	-2	-1	0	1	2	3	4	5	6	7	8	9	10
-3	0.0													
-2	0.0	0.0												
-1	0.0	0.0	0.0											
0	0.3	0.0	0.3	1.9										
1	1.3	0.0	0.4	14.7	36.2									
2	13.7	0.0	1.7	145.3	391.4	535.2								
3	142.1	0.0	5.3	243.8	840.1	2078.4	1537.8							
4	303.8	0.0	4.2	230.7	1200.8	2904.3	5548.1	3137.5						
5	471.6	0.0	2.2	166.5	945.1	2082.9	4373.7	9216.2	4222.3					
6	451.5	0.0	0.3	88.0	560.2	1162.4	3238.1	5653.1	6564.8	5492.9				
7	240.3	0.0	0.3	38.4	210.3	472.9	1226.7	3095.0	5616.0	6296.4	1888.7			
8	160.6	0.0	0.2	71.8	81.6	117.6	351.6	812.7	1783.1	2222.3	2048.2	520.9		
9	2.8		0.0	9.7	75.4	24.6	57.2	141.2	301.5	447.5	508.7	334.8	77.0	
10	0.2		0.0	0.8	6.3	30.9	16.4	18.2	29.4	47.4	54.5	56.8	29.1	29.1
11	0.0			0.1	0.3	0.2	3.9	17.2	7.2	7.7	7.6	9.3	17.9	20.8

times each ($pub\omega_1$, $pub\omega_2$) hit on the last 100m iterations, *1000. Everywhere, $pub\omega_1$>=$pub\omega_2$.

▓	above 5,000
▒	2,000-5,000
░	1,000-2,000

<u>Table 3</u>: What terminated the run: entry/exit/meeting.



	1 firm	-2	-1	0	1	2	3	4	5	6	7	8	9
0													
1													
2				░ 97%	░ 63%	▒ 86%							
3	░ 100%			░ 73%	▒ 65%	▒ 99%	▒ 100%						
4	▒ 52%			▓ 52%	▒ 75%	▒ 98%	▒ 100%	▒ 100%					
5	▒ 67%			▓ 67%	▒ 77%	▒ 98%	▒ 100%	▒ 100%	▒ 100%				
6	▒ 77%				▒ 62%	▒ 99%	▒ 100%	▒ 100%	▒ 100%	▒ 100%			
7	▒ 81%				▒ 60%	▒ 97%	▒ 100%	▒ 100%	▒ 100%	▒ 100%	▒ 100%		
8	░ 100%					▒ 94%	▒ 100%	▒ 100%	▒ 100%	▒ 100%	▒ 100%	▒ 100%	
9								▒ 100%	▒ 100%	▒ 100%	▒ 100%	▒ 100%	
10													

░	entry
▓	exit (without entry or call for meeting)
▒	call for meeting (without entry)

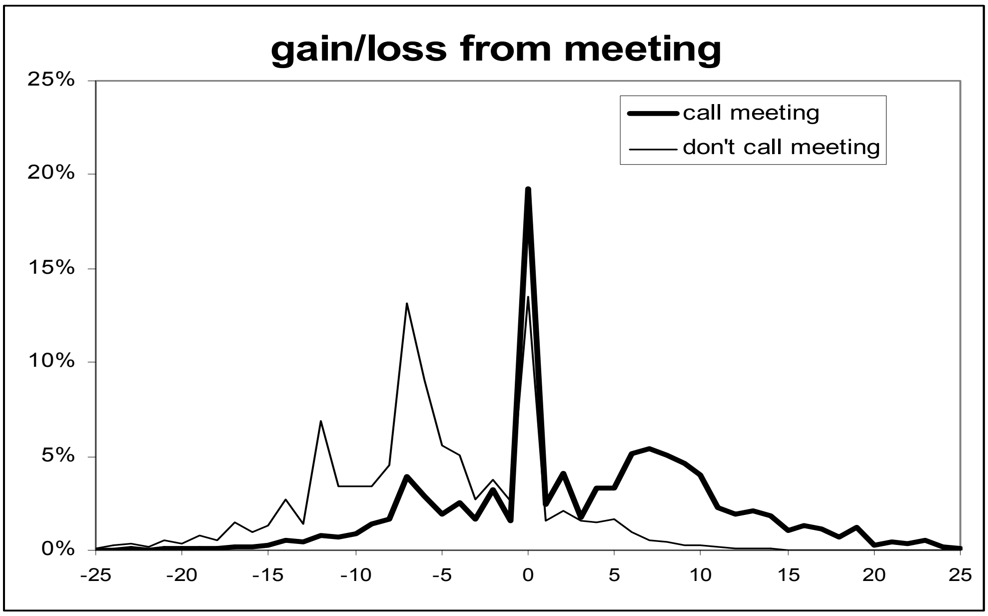
% - % of the runs terminated due to the dominant reason.

<u>Table 4</u>: Number of firms and the # periods between renegotiations.

Run length	renegotiation for any reason			meeting called without entry**		
	N=1	N=2	N=3	N=1	N=2	N=3
0	92.9%	3.3%	28.2%	96.9%	0.0%	0.2%
1	4.4%	38.6%	26.9%	2.1%	39.9%	34.1%
2	1.4%	34.6%	28.7%	0.6%	35.7%	42.6%
3	0.6%	15.9%	8.4%	0.2%	16.5%	11.5%
4	0.3%	6.1%	4.0%	0.1%	6.3%	5.8%
5	0.2%	0.5%	1.0%	0.0%	0.5%	1.5%
6	0.1%	0.2%	0.7%	0.0%	0.3%	1.1%
7	0.1%	0.3%	0.6%	0.0%	0.3%	1.0%
8	0.0%	0.2%	0.5%	0.0%	0.2%	0.7%
9	0.0%	0.1%	0.3%	0.0%	0.1%	0.4%
10+	0.0%	0.1%	0.7%	0.0%	0.1%	1.1%

** i.e., only the runs that were terminated by meeting without entry. Note that this includes costless monopoly renegotiation (also applies to N=2,3 if only 1 firm continues into the next period).

Figure 3: Distribution of gains/losses from the renegotiation.



Distribution of gains/losses from renegotiation, in terms of % of profit, all points (costly renegotiation without entry), profit gain/loss compared to the situation in which nobody renegotiates. Profit gain/loss are rounded to the nearest integer, thus e.g. 20% at zero is gain/loss from -0.5 to 0.5.

Table 5: Gains/Losses from price renegotiation.

	player who called meeting	**player who didn't call meeting**
% gained from meeting	**55.5%**	**11.6%**
% gained or no change from meeting*	**74.7%**	**25.1%**
% lost from meeting	**25.3%**	**74.9%**
top 1% gain	**+$27**	**+$10**
top 10% gain	**+14**	**+2**
worst 1% loss	**-16**	**-24**
worst 10% loss	**-7**	**-14**

* due to rounding to nearest integer, this actually includes loss up to -0.5.