Finite State Dynamic Games with Asymmetric Information: A Framework for Applied Work.

Chaim Fershtman and Ariel Pakes

September 22, 2008

Abstract

With applied work in mind, we define an equilibrium notion for dynamic games with assymetric information which does not require a specification for players' beliefs about their opponents types. This enables us to define equilibrium conditions which, at least in principal, are testable and can be computed using a simple reinforcement learning algorithm. We conclude with an example that endogenizes the maintenance decisions for electricity generators in a dynamic game among electric utilities in which the costs states of the generators are private information.

This paper develops a relatively simple framework for the applied analysis of dynamic games with asymmetric information. We consider a class of dynamic games in which there are a finite number of active players in each period, each characterized by a vector of state variables. Some of these state variables are publicly observed while others are private information. All state variables evolve over time with the outcome of the players' actions. In addition to affecting the evolution of the state variables these actions provide signals on the variables that are private information.

In this context we provide an equilibrium notion whose conditions are defined in terms of variables which, at least in principal, are observable. In particular they do not require a specification for players' beliefs about their opponents' types. This enables us to define equilibrium conditions which are testable, and the testing procedure does not require computation of posterior distributions. Moreover the equilibrium generated by a given set of primitives can be computed using a simple reinforcement learning algorithm. As a result one could view the AI algorithm as a description of how players' learn the implications of their actions in a changing environment and justify the output of the algorithm in that way.[1] Neither the iterative procedure which defines the algorithm, nor the test of the equilibrium conditions, are subject to a curse of dimensionality.

In the game we consider the number of active players may change over time due to entry and exit. Each active player is characterized by a vector of state variables (e.g. indexes of their cost function, qualities of the goods they market, etc.) which can take values only on a finite space. The state variables of one firm are not necessarily observable to the other firms. Each firm's returns in a given period are determined by the firms' state variables and their actions. The actions are allowed to contain a set of continuous controls (e.g. investment), and a set of discrete controls (e.g. sending a signal, entry, exit, etc.). The discrete control may or may not be observable to the firms' competitors and are taken from a finite set. The continuous controls are not observable to the firms' competitors and affects the game through their

[1] The reinforcement learning, or stochastic approximation, literature dates back to the classic paper of Robbins and Monroe,1956, and has been used extensively for calculating solutions to single agent dynamic programming problems (see Bertsekas and Tsikilis,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 asymmetric information dynamic games are much larger.

impact on the probability distribution of discrete state variables. In addition to these conditions we require a restriction on the informational structure of the game to guarantee that the state space be finite (there are alternative possibilities here, see below).

We define an Applied Markov Equilibrium (hereafter, AME) for our game as a triple which satisfies three conditions. The triple consists of; (i) a subset of the set of possible states, (ii) a vector of strategies defined for every possible information set of each agent, and (iii) a vector of values for every state that provides the expected discounted value of net cash flow conditional on the possible outcomes of the agent's actions at that state. The first condition is that the subset of states is a recurrent class of the Markov process generated by the equilibrium strategies. The second condition is that the strategies are optimal given the evaluations of outcomes for all points in this recurrent class, and the last condition is that these evaluations are indeed the expected discounted value of future net cash flows on the recurrent class if all agents play the equilibrium strategies.

When all the state variables are observed to all the agents, our equilibrium notion is similar to but weaker than the familiar notion of Markov Perfect equilibrium as used in Maskin and Tirole (1988, 2001). This because we only require that the evaluations be consistent with the outcomes of observed play on the recurrent class. More generally our equilibrium concept is closely related to the notion of Self Confirming equilibrium, as defined by Fudenberg and Levine (2001). Self Confirming equilibrium requires actions to be optimal given players' beliefs about opponents' actions and those beliefs to be correct on the equilibrium path. Our equilibrium conditions requires actions to be optimal given the players' evaluations of the outcomes of their actions and those evaluations to be consistent along the equilibrium path.

The fact that our equilibrium conditions are defined in terms of observables and hence can be tested, and that the equilibrium policies generated by a given set of primitives can be computed using a simple reinforcement learning algorithm, makes our notion of equilibrium relatively easy to use in applied work. To illustrate we conclude with an example that endogenizes the maintenance decisions of electricity generators. Maintenance decisions are naturally analyzed in a dynamic framework and can have large impacts on the performance of electricity markets. We take an admittedly simplified set of primitives and compute and compare AME equilibria in which the cost states of all generators are observable to an AME equilibria in which each firm only observes the cost states of its own generators. The AME equilibria are then comparred to the solution to a social planner and a monpolis'ts problem.

The next section describes the details of the game in a general setting. Section 2 provides a definition of, and sufficient conditions for, our notion of an Applied Markov Equilibrium. Section 3 provides an algorithm to compute this equilibrium, and section 4 contains our example.

1 A Finite State Dynamic Game with Asymmetric Information.

We extend the framework in Ericson and Pakes (1995) to allow for asymmetric information.[2] In each period there are n_t potentially active firms, and we assume that with probability one $n_t \leq \overline{n} < \infty$ (for every t). Each firm has payoff relevant characteristics. Typically these will be characteristics of the products marketed by the firm or of their cost functions. The profits of each firm in every period are determined by; the payoff relevant random variables of all of the firms, a subset of the actions (or controls) of all the firms, and a set of common determinants of demand and costs, say $d \in D$ where D is a finite set. For simplicity we assume that d_t is observable and evolves as an exogenous first order Markov process.

We make the following assumptions. The payoff relevant characteristics, which will be denoted by $\omega \in \Omega$, take values on a finite set of points. There will be two types of actions (or controls); actions which take values on a finite space, denoted by $m \in \mathcal{M}$, and actions which take values on a continuum, to be denoted by $x \in X$. It will be assumed that the continuous action of one firm is neither observed by the other firm nor a determinant of the profits of the other firm (this because we want to avoid signals which take values on a continuum). However the discrete actions of the firm are not restricted in either of these two ways. Both the continuous and discrete action can affect current profits and/or the probability distribution of payoff relevant random variables[3].

[2]Since we were motivated by our interest in dynamic oligopolies we will call our players "firms" and their payoffs as "profits".

[3]Note that these assumptions are similar to those used in the full information games considered by Ericson and Pakes (1995), and, as they do their, we could have derived the assumption that Ω is a finite set from more primitive conditions.

For notational simplicity we will assume that there is only one state variable, one discrete control, and one continuous control for each firm; i.e. that $\Omega \subset Z_+$, $\mathcal{M} \subset Z_+$, and $X \subset R$. In different models both the actions and the states will have different interpretations. Possibilities for actions include; maintenance and investment decisons, launching new products or sending a signal of the intention to launch, bidding in an auction and so on.

Letting i index firms, realized profits for firm i in period t are given by

$$\pi(\omega_{i,t}, \omega_{-i,t}, m_{i,t}, m_{-i,t}, x_{i,t}, d_t), \tag{1}$$

where $\pi(\cdot) : \Omega^n \times \mathcal{M}^n \times R \times D \rightarrow\ R$. Firms will be assumed to know their own (ω, x, m), but not necessarily their competitors (ω, m). We note that, in general, there may be a component of $\omega_{i,t}$ which has an impact on one firm's profits but not its competitors' profits (e.g. a component of $\omega_{i,t}$ may be the cost of x which varies across firms).

We assume that $\omega_{i,t}$ evolves over time with random firm specific outcomes, to be denoted by $\eta_{i,t}$, and a common shock that affects the ω's of all firms in a given period, say ν_t. Both η and ν take on values in a finite subset of $\mathcal{Z}_+$, say in $\Omega(\eta), \Omega(\nu)$ respectively. The transition rule is written as

$$\omega_{i,t+1} = F(\omega_{i,t}, \eta_{i,t+1}, \nu_{t+1}), \tag{2}$$

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

$$\mathcal{P}_\eta = \{\ P_\eta(.|\ m, x, \omega);\ \ m \in \mathcal{M},\ x \in X,\ \ \omega \in \Omega\}, \tag{3}$$

and is controlled by firm's choice of m and x, while the distribution of ν is given exogenously and equal to

$$\{p(\nu); \nu \in \Omega(\nu)\}.$$

Note that at least in this formulation of our problem we do not allow either the states or the actions of a firm's competitors to effect the evolution of the firm's own state variables.

The information set of each player at period t is, in principal, the history of variables that the player has observed up to that period. We restrict ourselves to a class of games in which strategies are a mapping from a subset of these variables, in particular to the variables that are observed and are either "payoff" or "informationally" relevant, where these two terms are defined as follows. The "payoff relevant" variables are defined, similar to Maskin and Tiorle (2001), to be those variables that are not current controls and affect the profits of at least one of the firms. In terms of equation (1), $(\omega_{i,t}, \omega_{-i,t}, d_t)$ will be payoff relevant. Observable variables that are not payoff relevant will be informationally relevant if and only if either; (i) even if no other agent's strategy depend upon the variable player i can improve its expected discounted value of net cash flows by conditioning on it (which implies that player i's optimal strategy depends on the variable), or (ii) even if player i's strategy does not condition on the variable there is at least one player j whose optimal strategy will depend on the variable. For example, say all players know $\omega_{j,t-1}$ but player i does not know $\omega_{j,t}$. Then even if player j does not condition its strategy on $\omega_{j,t-1}$, since $\omega_{j,t-1}$ can contain information on the distribution of the payoff relevant $\omega_{j,t}$, player i will generally be able to gain by conditioning its strategy on that variable.[4] As illustrated by the example, the variables that are informationally relevant at any point in time depend upon which of the payoff relevant variables are observed.

For simplicity we limit ourselves to the case where information is either known only to a single agent (it is "private"), or to all agents (it is "public"). Different models will allocate different states and actions to the publicly and privately observed components. The publicly observed vector will be denoted by $\xi_t \in \Omega(\xi)$, and the privately observed vector by $z_{i,t} \in \Omega(z)$. We will only consider games where both $\#\Omega(\xi)$ and $\#\Omega(z)$ are *finite*. We use the finiteness condition intensively in what follows and consider conditions on primitives which generate it in the next subsection.

If both decisions and the evolution of states conditional on those decisions depend only on $(\xi_t, z_{i,t})$, $(\xi_t, z_{i,t})$ evolves as a Markov process. More formally for all t

$$\xi_{t+1} = G_\xi(\xi_t, \nu_{t+1}, \epsilon_{t+1}), \tag{4}$$

where the distribution of ϵ_{t+1} is given by the family

$$\mathcal{P}_\epsilon = \{ \ P_\epsilon(\cdot| \ \xi, x, m, z, \eta, \ (\xi, x, m, \eta, z) \in \Omega(\xi) \times (X \times \mathcal{M} \times \Omega(\eta) \times \Omega(z))^n \ \}, \tag{5}$$

and hence can depend on $\{z_{i,t-1}\}$ (as it will if past information is revealed by current play). ϵ_{t+1} contains the payoff and informationally relevant information that is revealed over the current period and $G_\xi(\cdot)$ deletes public

[4]Note that these defintions will imply that an equilibrium in our restricted strategy space will also be an equilibrium in the general history dependent strategy space.

information which was either payoff or informationally relevant in t but no longer is in $t+1$.

Similarly

$$z_{i,t+1} = G_z(\xi_t, z_{i,t}, \mu_{i,t+1}), \tag{6}$$

and the distribution of $\mu_{i,t+1}$ is given by the family

$$\mathcal{P}_\mu = \{\ P_\mu(\cdot|\ \xi, x, m, z, \eta),\ (\xi, x, m, \eta, z) \in \Omega(\xi) \times (X \times \mathcal{M} \times \Omega(\eta) \times \Omega(z))^n\ \}. \tag{7}$$

Here $\mu_{i,t+1}$ contains the private information revealed to firm i over the current period and and $G_z(\cdot)$ deletes the private information which is no longer relevant.

Since the agent's information at the time actions are taken consists of $J_{i,t} = (\xi_t, z_{i,t}) \in \mathcal{J}_i$, we assume strategies are measurable $J_{i,t}$, i.e.

$$x(J_{i,t}):\ \ \mathcal{J}_i \rightarrow X, \ \text{ and } \ m(J_{i,t}):\ \ \mathcal{J}_i \rightarrow \mathcal{M}.$$

The timing of the game is as follows. At the beginning of each period there is a realization of $\{\mu, \nu, \epsilon, \}$. Firms then update their information sets with the updating functions (6) and (4). They then simultaneously decide on $\{m_{i,t}, x_{i,t}\}_{i=1}^{n_t}$. Finally we assume that firms maximize their expected discounted value of profits and have a common discount rate β , where $0 < \beta < 1$.

This formulation enables us to account for a range of institutional structures. The original Ericson and Pakes (1995) complete information assumptions is the special case where $\xi_t = (\omega_{i,t}, \omega_{-i,t})$ and $\epsilon_t = (\eta_{i,t}, \eta_{-i,t})$. When there is asymmetric information ϵ can depend on the actions of agents, as when signals are sent, and its contents can depend on both the private information of all agents in the preceding period (on $\omega \in \Omega^n$) and on the private information obtained in the current period (on $\eta \in \Omega(\eta)^n$). A simple example of a game with private information occurs when each firm knows the current value of its own ω but only last period's value of the ω's of its competitors (then $\xi_t = (\omega_{i,t-1}, \omega_{-i,t-1})$ and $\mu_{i,t} = \eta_{i,t}$, or equivalently $z_{i,t} = \omega_{i,t}$). These are the games considered in the recent econometric literature on dynamic games (see Pakes, Ostrovsky and Berry, 2007, and Bajari, Benkard and Levin, 2007)[5].

[5]They are particularly simple because in an Applied Markov Equilibrium, as defined below, these assumptions insure that agents never have to keep in memory more informa-

2 An Applied Markov Equilibrium.

Let s combine the information sets of all agents active in a particular period, that is $s = (J_1, \ldots, J_n)$ when each J_i has the same public component ξ. We will say that $J_i = (z_i, \xi)$ is a component of s if it contains the information set of one of the firms whose information is combined in s. Note also that we can write s more compactly as $s = (z_1, \ldots, z_n, \xi)$. So $\mathcal{S} = \{s : z \in \Omega(z)^n, \xi \in \Omega(\xi), \text{ for } 0 \leq n \leq \overline{n}\}$ lists the possible states of the world.

Any set of Markov strategies for all agents active at each $s \in \mathcal{S}$, together with an initial condition, defines a Markov process on $\mathcal{S}$. Recall that our assumptions insure that $\mathcal{S}$ is a finite set. As a result each possible sample path of this Markov process will, in finite time, wander into a recurrent subset of the states in $\mathcal{S}$, say $\mathcal{R} \subset \mathcal{S}$, and once in $\mathcal{R}$ will stay within it forever. That is though there may be more than one recurrent class associated with any set of policies, if a sample path enters a particular $\mathcal{R}$, a point, s, will be visited infinitely often if and only if $s \in \mathcal{R}$. Moreover the empirical distributions of visits to transitions in $\mathcal{R}$ will converge to a Markov transition kernel, say $p^{e,T} \equiv \{p^e(s'|s) : (s', s) \in \mathcal{R}^2\}$, while the empirical distribution of visits on $\mathcal{R}$ will converge to an invariant measure, say $p^{e,I} \equiv \{p^e(s) : s \in \mathcal{R}\}$. We let $p^e = (p^{e,T}, p^{e,I})$. It is understood that p^e is indexed by a set of policies and a particular choice of a recurrent class associated with those policies.

We now turn to our notion of Applied Markov Equilibrium. We build it from equilibrium conditions which could, at least in principle, be consistently tested. To obtain a consistent test of a condition at a point we must, at least potentially, observe that point infinitely often. So we limit ourselves to a definition of equilibrium that places conditions only at points that are in a recurrent class generated by that equilibrium.

As we shall see this weakens the traditional Nash conditions. In fact it generates equilibria which are closely related to the "self-confirming equilibria" introduced by Fudenberg and Levine (1993a), and share some of that equilibria's interpretive advantages[6]. On the other hand ours is probably the strongest notion of equilibrium that one might think could be empirically tested, as it assumes that the applied researcher doing the testing can access

tion than that contained in the prior period's ω and the current η_i. Note that whenever ω_{-i} has an independent effect on the profits of firm i (independent of other information known to the agent), and profits are privately observed, then μ_i at least contains (η_i, π_i).

[6]See also Dekel, Fudenberg and Levine (2004) for an anlysis of self confirming equilibrium in games with asymmetric information.

the union of the information sets available to the agents playing the game. We come back to these issues, and their relationship to empirical work, after we provide our definition of equilibrium.

Definition: Applied Markov Equilibrium. An applied Markov Equilibrium consists of

- A subset $\mathcal{R} \subset \mathcal{S}$;
- Strategies $(x^*(J_i), m^*(J_i))$ for every J_i which is a component of any $s \in \mathcal{S}$;
- Expected discounted value of current and future net cash flow conditional on realizations of η and a value for the discrete decision m, say $W(\eta, m|J_i)$, for each $(\eta, m) \in \Omega(\eta) \times \mathcal{M}$ and every J_i which is a component of any $s \in \mathcal{S}$,

such that

C1: $\mathcal{R}$ is a recurrent class. The Markov process generated by any initial condition $s_0 \in \mathcal{R}$, and the transition kernel generated by $\{(x^*, m^*)\}$, has $\mathcal{R}$ as a recurrent class (so, with probability one, any subgame starting from an $s \in \mathcal{R}$ will generate sample paths that are within $\mathcal{R}$ forever).

C2: Optimality of strategies on $\mathcal{R}$. For every J_i which is a component of an $s \in \mathcal{R}$, strategies are optimal given $W(\cdot)$, that is $(x^*(J_i), m^*(J_i))$ solve

$$max_{m \in \mathcal{M}} sup_{x \in X} \left[\sum_{\eta} W(\eta, m|J_i) p_{\eta}(\eta|x, m, \omega_i) \right],$$

and

C3: Consistency of values on $\mathcal{R}$. Take every J_i which is a component of an $s \in \mathcal{R}$. Let $\eta(x^*(J_i), m^*(J_i), \omega_i) \equiv \{\eta : p(\eta|x^*(J_i), m^*(J_i), \omega_i) > 0\}$. For every $\eta \in \eta(x^*(J_i), m^*(J_i), \omega_i)$

$$W(\eta, m^*(J_i)|J_i) = \pi^E\Big(J_i, m^*(J_i), x^*(J_i)\Big) \; +$$

$$\beta \sum_{J_i'} \Big\{ \sum_{\tilde{\eta}} W(\tilde{\eta}, m^*(J_i')|J_i') p(\tilde{\eta}|x^*(J_i'), m^*(J_i'), \omega_i') \Big\} p^e(J_i'|J_i, \eta),$$

and

$$\pi^E(J_i; m^*(J_i), x^*(J_i)) \equiv \sum_{J_{-i}} \pi_i \Big(\omega_i, m^*(J_i), x^*(J_i), \omega_{-i}, m^*(J_{-i}), d_t \Big) p^e(J_{-i}|J_i),$$

where

$$\Big\{ p^e(J_i'|J_i, \eta) \equiv \frac{p^e(J_i', \eta|J_i)}{p^e(\eta|J_i)} \Big\}_{J_i'}, \quad \text{and} \quad \Big\{ p^e(J_{-i}|J_i) \equiv \frac{p^e(J_{-i}, J_i)}{p^e(J_i)} \Big\}_{J_{-i}}. \spadesuit \quad (8)$$

Condition C2 states that at every J_i which is a component of an $s \in \mathcal{R}$ agent i chooses policies which are optimal with respect to the evaluations of outcomes determined by $\{W(\eta, m|J_i) : \eta \in \Omega(\eta), m \in \mathcal{M}\}$. Condition C3 states that at least for (η, m) combinations that have positive probability on the equilibrium path, these evaluations are the values that would be generated by p^e and the primitives of the problem if the agent played equilibrium strategies.

A few points are worth noting before moving on. First conditions C2 and C3 apply only to points in $\mathcal{R}$. In particular policies at points outside of $\mathcal{R}$ need not be optimal while the evaluations $\{W(\eta, m|J_i)\}$ need not be correct for J_i not a component of an $s \in \mathcal{R}$. Nor do we require consistency of the evaluations for the $W(\cdot)$'s associated with points in $\mathcal{R}$ but outcomes which have zero probability given equilibrium play[7].

Second none of our conditions are formulated in terms of beliefs about either the play or the "types" of opponents. There are two reasons for this to be appealing to the applied researcher. First, as beliefs are not observed, they can not be directly tested. Second, as we will show presently, it implies that we can compute equilibria without ever explicitly calculating posterior distributions.

[7]To see this last point note that the $W(\eta, m|J_i)$ for $\eta \notin \eta(x^*(J_i), m^*(J_i), J_i)$ or for $m \neq m^*$ are not required to satisfy C3. The only conditions on these evaluations are the conditions in C2; i.e. that chosing an $m \neq m^*$ and any x, or an x different x^* when $m = m^*$ would lead to a perceived evaluation which is less than that from the optimal policy. The fact that our conditions do not apply to points outside of $\mathcal{R}$ or to $\eta \notin \eta(x^*(J_i), m^*(J_i), J_i)$ implies that the conditional probabilities in equation (8) are well defined.

Comment 1. Our definition of Applied Markov Equilibria is closely related to the definition of Self Confirming Equilibria in Fudenburg and Levine (1993a). Self Confirming Equilibria weaken the standard Nash equilibrium conditions by requiring that each player's actions are optimal given the player's belief about opponent's actions but that these beliefs need only be correct along the equilibrium path (so no player observes actions which contradicts his beliefs). Our equilibrium conditions explicitly introduce the evaluations that the agents use to determine their optimal actions. These evaluations, together with the primitives of the problem, allow us to construct values for play along the equilibrium path. Our equilibrium condition insures that these values are consistent with optimizing behavior on points that are visited infinitely often. Of course data generated from a Self Confirming Equilibrium will also produce a set of evaluations. Moreover those evaluations will generate values that satisfy our AME conditions. However there is not a one to one relationship between the two concepts. This because the correct valuations could also be generated by incorrect perceptions on competitors' actions. In particular player i may have incorrect beliefs about the play of players j and k if those beliefs generate consistent values for the game. In addition our consistency requirement C3 is defined only for points in $\mathcal{R}$, and imposes no consistency conditions at points outside of $\mathcal{R}$ (though the game may start at points outside of $\mathcal{R}$). Self confirming equilibria requires that players have correct beliefs on opponents' actions also for points on the equilibrium path that are not in $\mathcal{R}$.

Comment 2. We now come back to the sense in which we can construct a consistent test of our equilibrium conditions. To determine what tests can be run we need to specify what information the empirical researcher has at its disposal. At best the empiricist will know the union of the information sets of all players at each period, that is our s_t. To determine what is testable when this is the case it will be useful to use a distinction introduced by Pakes and McGuire (2001). They partition the points in $\mathcal{R}$ into interior and boundary points. Points in $\mathcal{R}$ at which there are feasible (though inoptimal) strategies which can lead to a point outside of $\mathcal{R}$ are labelled boundary points. Interior points are points that can only transit to other points in $\mathcal{R}$ no matter which of the feasible policies are chosen (equilibrium or not). At interior $s \in \mathcal{R}$ we can obtain consistent estimates of $\{W(\eta, m|J_i), m \in \mathcal{M}, \eta \in \Omega(\eta)\}$ for every J_i which is a component of s. This plus the fact that the policies are observed implies that we can obtain a consistent test of C2 for all such points. However at an s which is a boundary point we only obtain consistent estimates of the $\{W(\cdot)\}$ for $(\eta, m) \in (\eta(x^*(J_i), m^*(J_i), \omega_i), m^*(J_i))$, that is for the states which are observed with positive probability given optimal strategies. This implies that we can only obtain a consistent test for a restricted notion of the optimality of strategies. In particular can not test against strategies that generate a positive probability of transiting to a point outside of the recurrent class. A test of the full set of our equilibrium conditions is provided in section 3.2 below.[8]

2.1 The Restriction to a Finite State Space.

In our definition of an AME we restricted ourselves to games where the state space was finite, and the next subsection provides an algoirthm which computes AME's for games with finite state spaces. We now breifly consider conditions which insure that the equilibrium we compute for a game in which we restrict the state space to be finite, will also be an equilibrium to the more general game in which strategies may be functions of the entire history of the game.

We have already assumed that there was: (i) an upper bound to the number of firms simultaneously active, (ii) each firm's physical states (our ω) could only take on a finite set of values, (iii) the discrete action was chosen from a finite feasible set, and (iv) the continuous action is not observed by the agent's opponents and affects the game only through its impact on the transition probabilities of the physical state. These restrictions would insure the condition that an equilibrium to a game with a finite state space is an equilbrium to a game with a more general state space were this a game of complete information (as in Ericson and Pakes, 1995)[9]. In our context

[8] Of course it is likely that the empiricist will observe less than s_t, perhaps only the publically available information in each period. Provided the empiricist knows (or has estimated) the primitive parameters, testing would then consist of computing the equilibrium associated with those primitives, and then testing whether the observed probabilities of transition from one public information set to another are consistent with the equilibrium calculations.

[9] They would also insure finiteness in a game with asymmetric information where the only source of asymmetric information is a firm specific state variable which is distributed independently over time (as in Bajari, Benkard and Levin, 2007, or Pakes Ostrovsky and Berry, 2007). In this case if the indepedently distributed private information took on an infinite range of values, then the strategies and values could take on an infinite range of

these restrictions insure that the payoff relevant random variables take values on a finite set of states, but they do not guarantee that there are a finite dimensional set of informationally relevant random variables; i.e. optimal strategies could depend on an infinite dimensional space of informationally relevant variabgles.

There are at least two possible ways to insure that the equilibrium we compute to our finite state game is an equilibrium to the unrestricted game. One is to consider a class of games in which the agents would not gain by distinguishing between more than a finite number of states. In general whether or not this is possible will depend on the functional forms and institutions relevant for a given situation. We provide one condition which insures it is possible presently. The second way to insure finiteness is to assume there are bounded cognitive abilities and these bounds generate a finite state space. The most obvious example would be a direct bound on memory (i.e. that agents can not remember what occured more than a finite number of periods prior), but there also could be bounds on complexity or perceptions that have a similar effect.

In our computational example we compute finite state equilibria generated by both types of assumptions. Indeed one of the questions we adress is whether the different ways we use to obtain finiteness, all three of which seem *a priori* reasonable, generate equilibria with noticeably different policies. The example of an institutional arrangement that insures that the equilibrium we compute for our finite state space is an equilibrium to the unrestricted game is one in which there is periodic simultaneous revelation of all variables which are payoff relevant to any agent.

Claim 1 . Periodic Revelation of Information. *If for any initial s_t there is a $T^* < \infty$ and a τ (whose distribution may depend on s_t) which is less than or equal to T^* with probability one, such that all payoff relevant random variables are revealed at $t - \tau$, then if we construct an equilibrium to a game whose strategies are restricted to not depend on information revealed more than τ periods prior to t, it is an equilibrium to a game in which strategies are unrestricted functions of the entire history of the game. Moreover there will be optimal strategies for this game which, with probability one, only take distinct values on a finite state space, so $\#|\mathcal{S}|$ is finite.* ♠

Sketch of Proof. Let $h_{i,t}$ denote the entire history of variables observed by

values, but the agents would only need to keep track of a finite set of continuation values.

agent i by time t, and $J_{i,t}$ denote that history truncated at the last point in time when all information was revealed. Let $(W^*(\cdot|J_i), x^*(J_i), m^*(J_i), p^e(\cdot|J_i))$ be AME strategies, valuations, and probability distributions when agents condition both their play and their expectations on J_i (i.e. they satisfy $C1, C2, C3$ above). Fix $J_i = J_{i,t}$. what we much show is that

$$(W^*(\cdot|J_{i,t}), x^*(J_{i,t}), m^*(J_{i,t}))$$

satisfy $C1, C2, C3$ if the agents' condition their expectations on $h_{i,t}$.

For this it suffices that if the '$*$' strategies are played then for every possible (J_i', J_{-i}),

$$p^e(J_i'|J_{i,t}, \eta) = Pr(J_i'|h_{i,t}, \eta), \text{ and } p^e(J_{-i}|J_{i,t}) = Pr(J_{-i}|h_{i,t}).$$

If this is the case strategies which satisfy the optimality comditions with respect to $\{W^*(\cdot|J_{i,t})\}$ will satisfy the the optimality comditions with respect to $\{W(\cdot|h_{i,t})\}$, where it is understood that the latter equal the expected discounted value of net cash flows conditional on all history.

We prove the second equality by induction (the proof of the first is similar and simpler). For the intial condition of the inductive argument use the period in which all information is revealed. Then $p^e(J_{-i}|J_i)$ puts probability one at $J_{-i} = J_{-i,t}$ as does $Pr(J_{-i}|h_i)$. For the inductive step, assume $Pr(J_{-i,t^*}|h_{i,t^*}) = p^e(J_{-i}|J_{i,t^*})$. What we must show is that if agents use the $*$ policies then the distribution of $J_{-i,t^*+1} = (\tilde{\mu}_{-i}, \tilde{\epsilon}, J_{-i,t^*})$ conditional on $h_{i,t^*+1} = (\tilde{\mu}_i, \tilde{\epsilon}, h_{i,t^*})$ depends only on $J_{i,t^*+1} = (\tilde{\mu}_{-i}, \tilde{\epsilon}, J_{i,t^*})$.

Use equations (6) and (5) to define

$$\tilde{\mu}_{-i} = G_z^{-1}(z_{-i,t^*+1}, z_{-i,t^*}), \ \tilde{\mu}_i = G_z^{-1}(z_{i,t^*+1}, z_{i,t^*}), \ \tilde{\epsilon} = G_\xi^{-1}(\xi_{t^*+1}, \xi_{t^*}),$$

and note that those assumptions imply that given the $*$ policies the distribution of $(\tilde{\mu}_{-i}, \tilde{\mu}_i, \tilde{\epsilon})$ conditional on (h_{i,t^*}, h_{-i,t^*}) depends only on (J_{i,t^*}, J_{-i,t^*}).

Since for any events (A, B, C), $Pr(A|B, C) = Pr(A, B|C)/Pr(B/C)$

$$Pr(J_{-i,t^*+1}|h_{i,t^*+1}) = \frac{Pr(\tilde{\mu}_{-i}, \tilde{\mu}_i, \tilde{\epsilon}, J_{-i,t^*}|h_{i,t^*})}{Pr(\tilde{\mu}_i, \tilde{\epsilon}|h_{i,t^*})}.$$

Looking first to the numerator of this expression, we have

$$Pr(\tilde{\mu}_{-i}, \tilde{\mu}_i, \tilde{\epsilon}, J_{-i,t^*}|h_{i,t^*}) = \sum_{J_{-i,t^*}} Pr(\tilde{\mu}_{-i}, \tilde{\mu}_i, \tilde{\epsilon}, J_{-i,t^*}|J_{i,t^*}, J_{-i,t^*})Pr(J_{-i,t^*}|h_{i,t^*}),$$

and from the hypothesis of the inductive arguement $Pr(J_{-i,t^*}|h_{i,t^*}) = p^e(J_{-i,t^*}|J_{i,t^*})$. A similar calculation for the denominator concludes the proof. ♠

3 An AI Algorithm to compute an AME.

In this section we show that we can construct an Applied Markov Equilibrium using a reinforcement learning algorithm. As a result our equilibria can be motivated as the outcome of a learning process.[10] In the reinforcement learning algorithm players have valuations regarding the continuation game and they choose their actions optimally given those valuations. Realizations of random variables whose distributions are determined by those actions are then used to update their evaluations. So in the algorithm players choose actions optimally given their evaluations, though their evaluations need not be correct. Note also that players are not engaged in intentional experimentation in the algorithm, however the algorithm can be designed to insure that many sample paths will be explored by providing sufficiently high initial valuations[11].

The algorithm provided in this section is iterative, and we begin by describing the iterative scheme. The rule for when to stop the iterations consists of a test of whether the equilibrium conditions defined above are satisfied, and we describe the test immediately after presenting the iterative scheme. We note that since our algorithm is a simple reinforcement learning algorithm, an alternative approach would have been to view the algorithm itself as the way players learn the values 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 testing subsection[12]. We conclude this section with a brief discussion of the properties of the algorithm; both its computational properties, and its relationship to various conceptual issues discussed in the economic literature.

[10]This is similar to in Fudenberg and Levine (1993a), but in our case the learning is about the value of alternative outcomes, while in their case it is about the actions of opponent players.

[11]This differs from Fudenberg and Kreps (1994) and Fudenberg and Levine (1993b) who considered models with active experimentation and studied the role of experimentation in the convergence of the learning process to a Nash equilibrium.

[12]On the other hand, there are several issues that arise were one to take the learning 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.

3.1 The Iterative Procedure.

Our algorithm approximates the $W \equiv \{W(\eta, m|J); \eta \in \Omega(\eta), m \in \mathcal{M}, J \in \mathcal{J}\}$ directly using techniques analogous to those used in the stochastic approximation (or reinforcement learning) literature (see footnote 3). The algorithm is iterative. An iteration, say k, is defined by couple

- its location, say $L^k = (J_1^k, \dots J_{n(k)}^k) \in \mathcal{S}$, defines the information set of the $n(k)$ agents active at iteration k[13], and
- a set of evaluations, W^k.

So to iterate we must update both L^k and the W^k.

Schematically the updates are done as follows. First the algorithm calculates policies for all agents active at L^k. These policies are chosen to maximize the agents' values (that is to solve condition C2) given the evaluations in memory, the W^k. Then computer generated random draws from the distributions which govern the innovations to both the public and private sources of information (from the distributions in equations (3), (5) and (7)) conditional on those policies and $L^k = (J_1^k, \dots J_{n(k)}^k)$ are taken. Those draws are used to update both L^k and W^k.

The location is updated using the updating functions in equations (4) (for the public information) and (6) (for the private information) for each of the active agents. This determines L^{k+1}. Next we update W^k. The k^{th} iteration only updates the components of W associated with L^k (it is *asynchronous*). It treats the updated $J_i^{k+1} = (\xi^{k+1}, z_i^{k+1})$ as a random draw on the next period's information set conditional on the chosen policies, and updates the W^k in memory at L^k with an average of the values at the updated state (as determined by the information in memory) and the initial values at L^k; a procedure which generates a W^{k+1} which is an average of the values obtained at L^k over all past iterations. We now formalize this procedure and then discuss some of its properties.

Details. The reinforcement learning part of the algorithm consists of an iterative procedure and subroutines for calculating initial values and profits. We begin with the iterative procedure. Each iteration starts with a location, L^k, and the objects in memory, say $M^k = \{M^k(J) : J \in \mathcal{J}\}$.

[13] Active agents include all incumbents, and in models with entry, the potential entrants.

Memory. 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
- $W^k(\eta, m|J)$ for $m \in \mathcal{M}$ and $\eta \in \Omega(\eta)$.

If $h^k(J) = 0$ there is nothing in memory at location J. If we require $W(\cdot|J)$ at a J at which $h^k(J) = 0$ we have an initiation procedure which sets $W^k(\eta, m|J_i) = W^0(\eta, m|J_i)$. We come back to a discussion of possible choices for W^0 below.

Policies and Random Draws for Iteration k. For each J_i^k which is a component of L^k call up $W^k(\cdot|J_i^k)$ from memory and choose $(x^k(J_i^k), m^k(J_i^k))$ to

$$max_{m \in M} sup_{x \in X} \left[\sum_{\eta} W^k(\eta, m|J_i^k) p_\eta(\eta|x, m, \omega_i^k) \right].$$

With this $\{x^k(J_i^k), m^k(J_i^k)\}$ use equation (1) to calculate the realization of profits for each active agent at iteration k[14]. These same policies, $\{x^k(J_i^k), m^k(J_i^k)\}$, are then substituted into the conditioning sets for the distributions of the innovations to the public and private information sets (the distributions in 3, 5 and 7), and they, in conjunction with the information in memory at L^k, determine a distribution for those innovations. A pseudo random number generator is then used to obtain a draw on those innovations, i.e. to draw $\Big((\eta_i^{k+1}, \mu_i^{k+1})_{i=1}^{n_k}, \epsilon^{k+1}, \nu^{k+1} \Big)$.

Updating. Use $\Big((\eta_i^{k+1}, \mu_i^{k+1})_{i=1}^{n_k}, \epsilon^{k+1}, \nu^{k+1} \Big)$ and the equations which determine the laws of motion of the public and private information (equations (4) and (6)) to obtain the updated location of the algorithm

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

[14]If d is random, then the algorithm has to take a random draw on it before calculating profits.

To update the W it is helpful to define a "perceived" value of play at iteration k after profits and the random draws are realized, i.e. to define

$$V^{k+1}(J_i^k) = \pi(\omega_i^k, \omega_{-i}^k, m_i^k, m_{-i}^k, x_i^k, d^k) \; + \tag{9}$$

$$max_{m \in M} sup_{x \in X} \; \beta \left[\sum_{\eta} W^k(\eta, m | J_i^{k+1}) p_{\eta}(\eta | x, m, \omega_i^{k+1}) \right].$$

Note that to calculate $V^{k+1}(J_i^k)$ we need to first find and call up the information in memory at locations $\{J_i^{k+1}\}_{i=1}^{n_{k+1}}$.[15] Once these locations are found we keep a pointer to them, as we will need to return to them in the next iteration.

For the intuition behind the update for $W^k(\cdot|J_i^k)$ note that were we to substitute the *equilibrium* $W^*(\cdot|J_i^{k+1})$ and $\pi^E(\cdot|J_i^k)$ for the $W^k(\cdot|J_i^{k+1})$ and $\pi^k(\cdot|J_i^k)$ in equation (9) above and use equilibrium policies to calculate expectations, then $W^*(\cdot|J_i^k)$ would be the expectation of $V^*(\cdot|J_i^k)$. Consequently we treat $V^{k+1}(J_i^k)$ as a random draw from the integral determining $W^*(\cdot|J_i^k)$ and update the value of $W^k(\cdot|J_i^k)$ as we do an average, i.e. we set

$$W^{k+1}(\eta, m|J_i^k) - W^k(\eta, m|J_i^k) = \frac{1}{A\Big(h^k(J_i^k)\Big)} [V^{k+1}(J_i^k) - W^k(\eta, m|J_i^k)], \tag{10}$$

where $A(\cdot) : \mathcal{Z}^+ \to Z^+$, is increasing, and satisfies Robbins and Monroe's conditions (1956)[16]. For example $A(h^k(J_i^k)) = h^k(J_i^k) + 1$, the number of times point J_i^k had been visited by iteration $k+1$, would satisfy those conditions and produces an estimate of $W^k(J_i^k)$ which is the simple average of the $V^r(J_i^r)$ over the iterations at which $J_i^r = J_i^k$. However since the early values of $V^r(\cdot)$ are typically estimated with more error than the later values, it is often useful to give them lesser weight. We come back to this point below.

Completing The Iteration. We now replace the $W^k(\cdot|J_k^i)$ in memory at location J_i^k with $W^{k+1}(\cdot|J_k^i)$ (for $i = 1, \ldots, n_k$) and use the pointers obtained

[15] The burden of the search for these states depends on how the memory is structured, and the efficiency of the alternative possiblities depend on the properties of the example. As a result we come back to this question when discussing the numerical example below.

[16] Those condition are that 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.

above to find the information stored in memory at L^{k+1}. This completes the iteration (we are now ready to compute policies for the next iteration). The iterative process is periodically stopped to run a test of whether the policies and values the algorithm outputs are equilbirium policies and values.

3.2 Testing For an Equilibrium.

This subsection assumes we have a W vector which is outputted at some iteration of the algorithm, say $W = \tilde{W}$, and provides a test of whether that vector generates AME policies and values on a recurrent subset of $\mathcal{S}$ determinned by $\tilde{W}$.

Once we substitute $\tilde{W}$ into condition C2 we determines policies for all agents active at each $s \in \mathcal{S}$. These policies determine the probabilities of transiting to any future state. Let the probability of transiting from s to s' be denoted by $q(s', s|\tilde{W})$, where $0 \leq q(s', s|\tilde{W}) \leq 1$, and $\sum_{s' \in \mathcal{S}} q(s', s|\tilde{W}) = 1$. Now order the states and arrange these probabilities into a row vector in that order, say $q(s|\tilde{W})$. Do this for each $s \in \mathcal{S}$, and combine the resultant rows into a matrix whose rows are ordered by the same order used to order the elements in each row. The result is a Markov matrix (or transition kernel) for the industry structures, say $Q(\cdot, \cdot|\tilde{W})$. This matrix defines the Markov process for industry structures generated by $\tilde{W}$. $Q(\cdot, \cdot|\tilde{W})$ is a finite state kernel and so any sample path generated by it and any initial condition will, with probability one in a finite number of periods, enter a recurrent class, say $\mathcal{R}(\tilde{W}, \cdot) \subset \mathcal{S}$, and once within it will stay within it forever. Our test consists of generating an $\mathcal{R}(\tilde{W}, \cdot) \subset \mathcal{S}$ and testing whether the $\{\tilde{W}(s);\ s \in \mathcal{R}(\tilde{W}, \cdot)\}$ satisfy the equilibrium conditions (C2 and C3 above).

To obtain our candidate for $\mathcal{R}(\tilde{W}, \cdot)$, we start at any s^0 and use $Q(\cdot, \cdot|\tilde{W})$ to simulate a sample path $\{s^j\}_{j=1}^{J_1+J_2}$. Let $\mathcal{R}(J_1, J_2, \cdot)$ be the set of states visited at least once between $j = J_1$ and $j = J_2$, and $\mathcal{P}(\mathcal{R}(J_1, J_2, \cdot))$ be the empirical measure of how many times each of these states was visited. Then, if we set $J_1 = J_1(J_2)$ and consider sequences in which both $J_1(J_2)$ and $J_2 - J_1(J_2) \to \infty$, $\mathcal{R}(J_2, \cdot) \equiv \mathcal{R}(J_1(J_2), J_2, \cdot)$ must converge to a recurrent class of the the process $Q(\cdot, \cdot|\tilde{W})$, and hence satisfies our condition C1. Call the recurrent subset of points obtained in this way $\mathcal{R}(\tilde{W})$. As we shall see it typically does not take long to generate a million iterations of the stochastic algorithm. As a result it is easy to simulate several million draws, throw out a few million, and then consider the locations visited by the remainder as the recurrent class.

Note that we have constructed $Q(\cdot,\cdot|\tilde{W})$ in a way that insures that condition $C2$ is satisfied everywhere. So what remains is to test whether condition $C3$ is satisfied at every $s \in \mathcal{R}(\tilde{W})$. One way to construct a test of this condition is already in the literature. Compute the integrals on the right hand side of the conditions defining the equilibrium W in C3 using the policies generated by $\tilde{W}$ to construct the probability distributions, $p^e(J_i'|J_i)$ and $p^e(J_{-i}|J_i)$, needed to compute the integrals. Then base the test statistic on the difference between the computed values and $\tilde{W}$. This is analogous to the test used by Pakes and McGuire (2001) and, as noted their, it is computationally burdensome. Even for moderately sized problems the computational burden of the test can greatly exceed the burden of the iterations leading to the test. To avoid this problem we now provide a test of condition C3 which does not require explicit computation of the integrals on the right hand side of that condition, and has a transparent interpretation as a measure of the extent of approximation error in our estimates.

Instead of computing the right hand side of the integrals in condition C3 directly the test approximates them using simulation and then accounts for the simulation error in the approximations. That is our test consists of obtaining the differences between the estimates of $\tilde{W}(\eta, m^*(\tilde{W})|J_i)$ in memory, and an approximation to the expected discounted values of future net cash flows that an agent with the information set J_i and the random draw η_i would obtain were all agents using the policies generated by $\tilde{W}$. The approximation is a sample average of the discounted value of net cash flows from simulated sample paths starting at (J_i, η_i). The squared differences between the $\tilde{W}$ and the average of the discounted value over the simulated sample paths is a sum of; (i) the sampling variance in the average of the discounted value of the simulated sample paths (we will refer to this as the sampling variance term), and (ii) the difference between the *expectation* of the discounted net cash flows from the simulated paths and the relevant components of $\tilde{W}$ (we will refer to this as the "bias" term). We subtract a consistent estimate of the sampling variance term from this squared difference to obtain a test statistic which, at least in the limit, will depend only on the bias term and have an interpretation as the percentage bias in our estimates.

Details. The test is constructed as follows. Start at an initial $s^0 \in \mathcal{R}$ and an initial draw on η for each J_i component of s^0, i.e. at a set of couples $\{(J_i^0, \eta_i^0)\}_{i=1}^{n^0}$, where n^0 is the number of active agents at s^0. Now simulate draws for $\Big((\eta_i^1, \mu_i^1)_{i=1}^{n^0}, \epsilon^1\Big)$ using the policies generated by $\tilde{W}$. Use these simulation draws to compute

$$\hat{W}^{l=0}(\eta, m^*(J_i^0)|J_i^0) \equiv \pi(J_i^0, J_{-i}^0, m^*(J_i^0), m^*(J_{-i}^0), x^*(J_i^0), d^0) + \beta\tilde{W}(\eta_i^1, m^*(J_i^1)|J_i^1),$$

where it is understood that $m^*(\cdot)$ provides the policies generated by $\tilde{W}$, and

$$J_i^1 = (\xi^1, z_i^1), \quad \xi^1 = G_\xi(\xi^0, \nu^0, \epsilon^1), \quad z_i^1 = G_z(\xi^0, z_i^0, \mu_i^1),$$

for each of the n^0 points (J_i^0, η_i^0).

Then, for $i = 1, \ldots, n$, keep in memory at a location which is (J_i, η_i) specific: (i), $\hat{W}^0(\eta, m^*(J_i^0)|J_i^0)$, (ii), the square of this, or $S\hat{W}^0(\eta, m^*(J_i^0)|J_i^0) = \hat{W}^0(\eta, m^*(J_i^0)|J_i^0)^2$, and (iii) an initialized counter, say $h^{l=0}(J_i, \eta_i) = 1$.

Now consider the simulated locations $\{(J_i^1, \eta_i^1)\}_{i=1}^{n^1}$. At each of these points simulate as above and compute

$$\hat{W}^1(\eta, m^*(J_i^1)|J_i^1) \equiv \pi(J_i^1, J_{-i}^1, m^*(J_i^1), m^*(J_{-i}^1), x^*(J_i^1), d^1) + \beta\tilde{W}(\eta_i^2, m^*(J_i^2)|J_i^2).$$

If (J_i^1, η_i^1) is the same as one of the values (J_i^0, η_i^0), average the two values of $\hat{W}(\cdot)$ and $S\hat{W}(\cdot)$ at that location, call the averages $A\hat{W}^1(\cdot)$ and $AS\hat{W}^1(\cdot)$, and keep them together with a value for $h^l(\cdot)$ equal to 2 in memory at that location. If a particular (J_i^1, η_i^1) was not visited prior to this start a new location, setting $\hat{W}^1(\cdot|\cdot)$, $S\hat{W}^1(\cdot)$, and $h^1(J_i^1, \eta_i^1)$ as above. We continue in this manner until a large number of periods are simulated.

If we let E take expectations over the simulated random draws then

$$E\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)} - 1\Big)^2 = E\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i) - E\Big[A\hat{W}(\eta_i, m^*(J_i)|J_i)\Big]}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big)^2 + \Big(\frac{E\Big[A\hat{W}(\eta_i, m^*(J_i)|J_i)\Big] - \tilde{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big)^2. \quad (11)$$

The first term after the equality in (11) is the sampling variance, while the second term is the bias, both expressed as a fraction of the evaluations outputted by the program.

Moreover if we let

$$\hat{Var}\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big) \equiv \quad (12)$$

$$\frac{A\hat{SW}(\eta_i, m^*(J_i)|J_i)\frac{h(\eta_i,J_i)}{h(\eta_i,J_i)-1} \; - \; A\hat{W}(\eta_i, m^*(J_i)|J_i)^2}{\tilde{W}(\eta_i, m^*(J_i)|J_i)^2},$$

then

$$E\Big[V\hat{a}r\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big)\Big] = E\Bigg(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i) - E\Big[A\hat{W}(\eta_i, m^*(J_i)|J_i)\Big]}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Bigg)^2,$$

and we have an unbiased estimate of the sampling variance. Consequently if

$$B\hat{i}as(AW(\eta_i, m^*(J_i)|J_i))^2 \equiv \tag{13}$$

$$\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)} - 1\Big)^2 - V\hat{a}r\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big),$$

then $B\hat{i}as(AW(\eta_i, m^*(J_i)|J_i))^2$ is an unbiased estimate of the square of the percentage bias in $A\hat{W}(\eta_i, m^*(J_i)|J_i))$. Since higher order moments of this estimate are finite, any weighted average of independent estimates of the bias terms over the recurrent class of points will converge to the same weighted average of the true bias term across these points (a.s.).

Let $\mathcal{P}_{\mathcal{R}(\tilde{W})}(s)$ provides the fraction of times point $s \in \mathcal{R}(\tilde{W}) \subset \mathcal{S}$ is visited in constructing $\mathcal{R}$ (i.e. visited between iterations $J_1(J_2)$ and J_2 in that construction), and n_s be the number of agents active at s (assuming the policies generated by $\tilde{W}$). Then our test statistic, to be denoted by $\mathcal{T}$, is an $L^2(\mathcal{P}_{\mathcal{R}(\tilde{W})})$ norm in the bias terms defined in equation (13). More formally

$$\mathcal{T}(\cdot) \equiv \Bigg\| n_s^{-1} \sum_{i=1}^{n_s} \sum_{\eta_i} \frac{h(\eta_i, J_i)}{\sum_{\eta_i} h(\eta_i, J_i)} B\hat{i}as(AW(\eta_i, m^*(J_i)|J_i))^2 \Bigg\|_{L^2(\mathcal{P}_{\mathcal{R}(\tilde{W})})}.$$

Assuming the computer's calculations are exact, $\mathcal{T}$ will tend to zero as the number of simulation draws used in the test grows large if and only if $\tilde{W}$ satisfies condition C3. More generally $\mathcal{T}$ is a consistent estimate of the average percentage difference between the two sides of that fixed point in C3. We assume we are "at an equilibrium" when it is sufficiently small[17].

[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)

3.3 Properties of the Algorithm.

Recall that our equilibrium conditions do not require us to form beliefs about player's *types*. Analgously our algorithm does not require us either to compute such beliefs or test for their consistency with the actual distribution of types. We are able to do this by basing our equilibrium concept on the consistency of the players' evaluations of the possible outcomes of their behavior, and then using a stochastic algorithm to estimate those evaluations.

The advantages of using a stochastic algorithm to compute the recurrent class of equilibria in full information games were explored by Pakes and McGuire (2001).[18] They note that, at least formally, the stochastic algorithm they propose does away with all aspects of the curse of dimensionality in computing equilibria except for the computation of the actual test statistic. On the other for the games that they analyze the computation of the test statistic quickly becomes the dominant computational burden. We circumvent this problem by substituting simulation for explicit integration in the construction of the test statistic, so there is no necessary curse of dimensionality in our algorithm.

However as is typical in algorithms designed to compute equilibria for (nonzero sum) dynamic games, there is no guarantee that our algorithm will converge to equilibrium values and policies; that is all we can do is test whether the algorithm outputs equilibrium values, we can never guarantee convergence to an equilibrium *a priori*. Moreover there may be more than one equilibria which is consistent with a given set of primitives, in which case

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. before we proceeded in this way we would want to formalize tradeoff between size, power, and the number of simulation draws.

[18]Were we to consider a computational comparison of our stochastic algorithm to an algorithm designed to implement an asymmetric information equilibrium concept that required the computation of consistent posteriors, the computational advanatages of the stochastic algorithm would be even greater. If we were to use constructed posteriors the expectation required to update continuation values would involve a convolution over the distributions induced by the competitors' policies at different possible states. This, in turn, would 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. In rather stark contrast, the updating burden of the stochastic algorithm remains the same; it still only need to update averages.

the way we initiate the algorithm, i.e. our choice for W^0, and our updating procedure will implicitly select out the equilibrium computed. High initial values are likely to encourage experimentation, and lead to an equilbria in which players have explored many alternatives. On the other hand, precisely for the same reason, high initial values will tend to result in a longer computational times and a need for more memory.

There are other aspects of the algorithm that can be varied as well. Our test insures that the $\tilde{W}$ outputted by the algorithm is consistent with the distribution of current profits and the discounted evaluations of the next period's state. We could have considered a test based on the distribution of discounted profits over τ periods and the discounted evaluation of states reached in the τ^{th} period. We chose $\tau = 1$ because it generates the stochastic analogue of the test traditionally used in iterative procedures to determine whether we have converged to a fixed point. It may well be that a different τ provides a more discerning test, and with our testing algorithm it is not computational burdensome to increase τ.

Finally since our estimates of the $\tilde{W}$ are formed as sample averages, we expect the estimates from a particular location to be more accurate the more times we visit that location (the larger $h(\cdot)$). If one is particularly interested in policies and values at a given point, for example at a point that is consistent with the current data on a given industry, one can increase the accuracy of the relevant estimates by restarting the algorithm repeatedly at that point.

4 Example: Maintenance Decisions in An Electricity Market.

The restructuring of electricity markets has focused attention on the design of markets for electricity generation. One issue in this literature is whether the design would allow generators to make super-normal profits during periods of high demand. This because of the twin facts that currently electricity is not storable and has extremely inelastic demand (for a review of the literature on price hikes and a careful empirical analysis of their sources during the California price increaes of the summer of 2000, see Borenstein, Bushnell, and Wolak, 2002). The analysis of the sources of price increases during periods of high demand typically conditions on whether or not generators are bid into or withheld from the market, though some of the literature have tried to incorporate the possiblity of "forced" (in constrast to "scheduled") outages (see Borenstein, et.al, 2002). Scheduled outages are largely for maintenance and maintenance decisions, as many authors have noted, are endogenous. This line of reasoning has led to an extensive empirical literature on when firms do maintenance (see, for e.g. Harvey, Hogan and Scatzki, 2004, and the literature reviewed their).[19]

Since the benefits from incuring maintenance costs today depend on the returns from bidding the generator in the future, and the latter depend on what the firms' competitors bid at future dates, a consistent framework for analyzing maintenance decisions requires dynamics with strategic interaction. To the best of our knowledge maintenance decisions of electric utilities have not been analyzed within such a framework to date. Here we provide a simple model that endogenizes maintenance decisions, and then ask how asymmetric information effects the results.

Overview of the Model. In our base model the level of costs of a generator evolve on a discrete space in a non-decreasing random way until a maintenance decision is made. In the full information model each firm knows the current cost state of its own generators as well as those of its competitors. In the asymmetric information the firm knows the cost position of its own generators, but not of those of its competitors.

Firms can hold their generators off the market for a single period and do maintenance. Whether they do or do not do maintenance is public information. If they do maintenance the cost level of the generator reverts to a base state (to be designated as the zero state). If they do not do maintenance they bid a supply function and compete in the market. If a generator is operated in the period its costs increase stochastically. There is a regulatory rule which insures that the firms do maintenance on each of their generators at least once every six periods.

For simplicity we assume that if a firm submits a bid function for producing electricity from a given generator, it always submits the same function (so in the asymmetric information environment the only cost signals sent by the firm is whether it does maintenance on each of its generators). We do, however, allow for heterogeneity in both cost and bidding functions across generators. In particular we allow for one firm which owns only big genera-

[19]For an empirical investigation of the co-ordination of maintenance decsions, see Patrick and Wolak, 1997.

tors, Firm B, and one firm which only owns small generators, Firm S. Doing maintenance on a large generator and then starting it up is more costly than doing maintenance on a small generator and starting it up, but once operating the large generator operates at a lower marginal cost. The demand function facing the industry distinguishes between the five days of the work week and the two day weekend, with demand higher in the work week.

In the full information case the firm's strategy are a function of; the cost positions of its own generators, those of its competitors, and the day of the week. In the asymmetric information case the firm does not know the cost position of its competitor's generators, though it does realize that its competitors' strategy will depend on those costs. As a result any variable which helps predict the costs of a competitors' generators will be informationally relevant.

In the asymmetric information model Firm B's perceptions of the cost states of Firm S's generators will depend on the last time each of Firm S's generators did maintenance. So the time of the last maintenance decision on each of Firm S's generators are informationally relevant for Firm B. Firm S's last maintenance decisions depended on what it thought Firm B's cost states were at the time those maintenance decisions were made. Consequently Firm B's last maintenance decisions will generally be informationally relevant for itself. As noted in the theory section, without further restrictions this recurrence relationship between one firm's actions at a point in time and the prior actions of the firm's competitors at that time can make the entire past history of maintenance decisions of both firms informationally relevant. Below we consider three separate restrictions each of which have the effect of truncating the relevant past history in a different, and we think reasonable way. We compute an AME for each one of them, and then compare all results.

4.1 Details and Parameterization of The Model.

Firm B has three generators at its disposal. Each of them can produce 25 megawatts of electricity at a constant marginal cost which depends on their cost state (ω). Firm S has four generators at its disposal each of which can produce 15 megawatts of electricity at a constant cost which depends on their cost state. Firm B's generator's marginal cost is smaller than those of Firm S at any cost state, but the cost of maintaining and restarting firm B's generators is two and a half times that of maintaining and restarting Firm

Table 1: **Primitives Which Differ Among Firms.**

Parameter	Firm B	Firm S
Number of Generators	3	4
Range of ω	0-4	0-4
Generator Capacity	25	15
Marginal Cost ($\omega = (0, 1, 2, 3)$)*	(20,60,80,100)	(50,100,150,200)
Costs of Maintenance	15,000	6,000

* At $\omega = 4$ the generator must shut down.

S's generators (see table 1).

The firms bid just prior to the production period and they know the cost of their own generators before they bid. If a generator is bid it bids a supply curve which is horizontal at the highest marginal cost at which it can operate (a hundred dollars per megawatt hour for the big generators and two hundred dollars for the small) until its maximum capacity; the curve then turns vertical. The market maker runs a uniform price auction. The bids are horizontally summed and the resultant supply curve is intersected with the demand curve to determine the price per megawatt hour. This price is paid for each megawatt hour accepted by the market maker.

The demand curve is log-linear

$$log(Q) = D_d - \alpha log(P),$$

with a price elasticity of $\alpha = .3$ and a level which is about a third higher on weekdays than weekends (i.e. $D_{d=weekday} = 8.5, D_{d=weekend} = 6.5$).

If the generators bid is accepted, the generator is operated and the state of the generator stochastically decays. Formally if $\omega_{i,j,t} \in \Omega = \{0, 1, \ldots, 4\}$ is the cost state of firm i's j^{th} generator in period t, then

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

where, if the generator is operated in the period

$$\eta_{i,j,t} = \begin{cases} 0 & \text{with probability } .1 \\ 1 & \text{with probability } .4 \\ 2 & \text{with probability } .5. \end{cases}$$

If, on the other hand, the generator is not operated in this period it does maintenance and at the begining of the next period is ready to be operated at the low cost base state ($\omega = 0$).

The information at the firm's disposal when it makes its maintenance decision, say $J_{i,t}$, always includes the vector of states of its own generators, say $\omega_{i,t} = \{\omega_{i,j,t}; j = 1 \ldots n_i\} \in \Omega^{n_i}$, and the day of the week (denoted by $d \in D$). In the full information it also includes the cost states of its competitors' generators and the strategies are functions of these variables. In the asymmetric information case firms' do not know their competitors' cost states and so keep in memory public information sources which may help them predict their competitors' actions. The specification for the public information used in this excercise differs for the different asymmetric information models we run, so we come back to it when we introduce those models.

The strategy of firm i is a choice of

$$m_i = [m_{1,i}, \ldots m_{n_i,i}] : \mathcal{J} \rightarrow \Pi_j \Big(0, m_i\Big) \equiv M_i,$$

where m_i is the bid function. As noted this always consists of two numbers; the highest marginal cost at which it operates its generators, and its capacity. We assume that whenever the firm withholds a generator from the market they do maintenance on that generator, and that maintenance must be done at least once every six periods.[20] The cost of that maintenance is denoted bu cm_i.

The profit function is given by $\pi_i : M_S \times M_B \times \Omega^{n_i} \times D \rightarrow \mathcal{R}_+$ and is defined as

$$\pi_i\Big(m_{B,t}, m_{S,t}, d_t, \omega_{i,t}\Big) = p(m_{B,t}, m_{S,t}, d_t) \sum_j y_{i,j,t}(m_{B,t}, m_{S,t}, d_t)$$

$$-\sum_j \Big[I\{m_{i,j,t} > 0\} c(\omega_{i,j,t}, y_{i,j,t}(m_{B,t}, m_{S,t}, d_t)) - I\{m_{i,j,t} = 0\} cm_{j,i}\Big],$$

where $p(m_{1,t}, m_{2,t}, d_t)$ is the market clearing price, $y_{i,j,t}(m_{B,t}, m_{S,t}, d_t)$ is the output alocated by the market maker to the j^{th} generator of firm i, $I\{\cdot\}$ is the indicator function which is one if the condition inside the brackets is satisfied and zero elsewhere, and $c(\omega_{i,j,t}, y_{i,j,t}(\cdot))$ is the cost of producting output $y_{i,j,t}$ at a generator whose cost state is given by $\omega_{i,j,t}$.

[20] In none of our runs was this constraint binding more than in .29% of the cases, and in most cases it never bound at all.

Note. We now go on to describe the different sources of public information that we allow the firm to condition its expectations, and hence its strategies, on in the three asymmetric information models whose equilibria we compute. We want to point out, however, that all three are quite simple special cases of our general model. In particular none of them allow for

- either a continuous control or entry and exit, or for
- ω_{-i} to enter the profits of firm i.

These simplifications make the example particularly easy to compute and its results easy to interpret since they imply that: (i) the only additional information accumulated over a period on the likely actions of the firm's competitors is m_{-i}, and (ii) the only response to that information that we have to focus on is m_i. We want to point out, however, that they simplifications are neither necessary given our setup, nor are they likely to generate an adequate approximation to any real electricity market. They were chosen to make it easier for us to isolate the impact of asymmetric information on equilbrium behavior.

4.2 Alternative Informational Assumptions for the Assymetric Information Model.

As noted the public information that is informationally relevant in the sense that it helps predict the maintenance decisions of the firm's competitor could, in principal, include all past maintenance decisions of all generators; those owned by the firm as well as those owned by the firms' competitors. In order to apply our framework we have to insure that the state space is finite. We present results from three different "natural" assumptions each of which have the effect of insuring finiteness and compare their computational properties.

All three asymmetric information (henceforth, AI) models that we compute are based on exactly the same primitives and assume $(\omega_{i,t}, d_t) \in J_{i,t}$. The only factor that differentiates the three is the public information kept in memory to help the firm assess the likely outcomes of its actions. Two of the alternatives assume bounded recall; in one a firm partitions the history it does remember more finely than in the other. The third case is a case of periodic full revelation of information. This case assumes there is a regulator who inspects all generators during every fifth period and announces the states of all generators just before the sixth period.

The public information kept in memory in the three asymmetric information models is as follows.

1. In finite history $''\tau''$ the public information is the time since the last maintenance decision of each generator (recall that since all generators must do maintenance at least once every six periods, so $\tau \leq 5$).

2. In finite history $''m''$ the public information is the maintenance decisions made in each of the last five periods on each generator.

3. In the model with periodic full revelation of information the public information is the state of all generators at the last date information was revealed, and the maintenance decisions of all generators since that date (reall that full revelation occurs every sixth period, so no more than five periods of maintenance decisions are ever kept in memory).

The information kept in memory in each period in the first model is a function of that in the second; so a comparison of the results from these two models provides an indication on whether the extra information kept in memory in the second model has any impact on behavior. The third model, the model with full revelation every six periods, is the only model whose equilibrium is insured to be an equilibrium to the game where agents can condition their actions on the indefinite past. I.e. there may be unexploited profit opportunties when employing the equilibrium strategies of the first two models. On the other hand the cardinality of the state space in the model with full revelation of information is an order of magnitude larger than in either of the other two models.[21]

4.3 Reference Models: The Social Planner and Monopoly Solutons.

To evaluate the performance of the AI and full information (to be labelled FI) models we compare their results to those that we would obtain from the same primitives were maintenance decisions made by:

[21] This does not imply that at every instant the memory requirements of one are greater than the other, or that the recurrent class of one is larger than the other. So the fact that the model with full revelation has a much larger state space does not imply that it has larger memory requirements. The size of memory implications and computational burden of the different assumptions have to be analyzed numerically.

- A social planner with full information who maximizes the sum of the discounted value of consumer surplus and net cash flows to the firms.
- A monopolist with full information which choses maintenance to maximize the sum of net cash flows to the two firms.

The social planner provides us with the efficient allocation of maintenance times. The monopolist provides us with the allocation of maintenance times that maximize the sum of the firms' discounted returns. The monopoly and social planner problems are single agent problems with (generically) unique optimal policies. They were computed using a standard contraction mapping[22].

4.4 Computational Details and Results.

The AME equilibrium for each of our four duopolies was computed using the algorithm provided in section 3. This section describes the model-specific details needed for the computation and provides computational properties of the results. The details include; (i) starting values for the $W(\cdot|\cdot)$'s and the $\pi^E(\cdot|\cdot)$, (ii) information storage procedures, and (iii) the testing procedure. The computational properties include; (i) test results (ii) compute times, and (iii) sizes of the recurrent class.

To insure experimentation with alternative strategies we used starting values which were likely to be much higher than their true equilibrium values. In particular our intitial values for expected profits are the actual profits the agent would receive were its competitor not bidding at all, or

$$\pi_i^{E,k=0}(m_i, J_i) = \pi_i(m_i, m_{-i} = 0, d, \omega_i).$$

For the intial condition for the expected discounted values of outcomes (η_i) for different strategies we assumed that the profits were the other competitor not producing at all could be obtained forever with zero maintenance costs, that is

$$W^{k=0}(\eta_i, m_i|J_i) = \frac{\pi_i(m_i, m_{-i} = 0, d, \omega_i + \eta_i(m_i))}{1 - \beta}.$$

[22]The equilibrium concept for the full information duopoly is a special case of that for the game that allows for asymmetric information (it corresponds to the equilibrium concept used in Pakes and McGuire, 2001). It was computed using the same techniques as those used for the AI duopoly (see section 3 and the details we now turn to).

The memory was structured first by public information, and then for each given public information node, by the private information of each agent. We used a tree structure to order the public information and a hash table to allocate the private information conditional on the public information. To keep the memory manageable, every fifty million iterations we performed a "clean up" operation which dropped all those point which were not visited at all in the last ten million iterations.

The algorithm was set up to perform the test every one hundred million iterations. Recall that the test statistic is an $\mathcal{L}^2(P)$ norm in the percentage deviation between the simulated and estimated values normalized by the variance in simulated value (where P is determined by the frequency of points visited); roughly an R^2 for the fit of the simulated to the estimated values. Had we used the test to determine the stopping iteration and stopped the algorithm whenever the R^2 from our test was above .995, we would have always stopped the test at either 100 or 200 million iterations.

Since we wanted more detail on how the test statistic behaved at a higher number of iterations we ran each of our runs for one billion iterations. There was no perceptible change in the test statistic after the 300 millionth iteration. To illustrate how the test behaved we computed one run of the full revelation model that stopped to do the test every ten million iterations. Figure 1 graphs the results from those tests. As can be seen from that figure the R^2 increased rapidly until about 100 million iterations. It increases fruther but at a less rapid rate between 100 and 130 iterations, and remains essentially unchanged after 150 million iterations (at a value of about .9975; see figure 1). We could have decrased the number of iterations significantly were we willling to use starting values that were not as high.

The time per one hundred million iterations, each of which includes the test time, is reported in table 2.[23] The differences in compute times across models roughly reflect the differences in the size of the recurrent class from the different specifications, as this determines the search time required to bring up and store information.

There are some notable differences in the sizes of the recurrent class across models. First the recurrent class in the finite history τ model is less than half the size of those in the other AI models. Second, though the cardinality of

[23] All computations were done using a Linux Red-Hat version 3.4.6-2 operating system. The machine we used had seven AMD Opteron(tm) processors 870; CPU: 1804.35 MHz, and 32 GB RAM.

Table 2: **Computational Comparisons.**

	AI; Finite Hist. τ	AI; Finite Hist. m	AI; Full Revel.	Full Info.
Compute Times per 100 Million Iterations (Includes Test).				
Hours	1.05	2.37	2.42	2.44
Cardinality of Recurrent Class.				
Firm B ($\times 10^{-6}$)	.349	.808	.990	.963
Firm S ($\times 10^{-6}$)	.447	.927	1.01	1.09

the state space for the AI model with periodic full revelation of information is an order of magnitude larger than in any of the other models, there is very little difference between the size of its recurrent class and either the recurrent class of the finite history m AI model or the FI model. Note that this implies that if we limit our attention to the recurrent classes of models, both the computational burden and the memory requirements from the AME AI model are similar to those from the FI model.

After computing policies we ran a one million iteration simulation using the computed policies for each of our models. We now turn to the numerical results from these runs. Table 3 provides a comparison of the results from from the three AI models. It is suprising how little difference there is in these statistics across the three models. This may well be a result of the particular parameterization that underlies our computation and in the way we implement our bounded memory assumptions. Whatever the reason, the remainder of the paper focuses on the results for the AI model in which there is periodic full revelation of information.

4.5 Numerical Results.

The output of the algorithm includes a recurrent class of states as well as strategies, costs (both operational and maintenance), and profits at those states. Here we focus on the maintenance decisions and their implications on consumer welfare and profits. For this we use the output of the simulation runs we initiated after computing the equilibrium. We begin with the social planner and monopoly problems as they provide clear reference points.

Table 3: **Three Asymmetric Information Models.**

	Finite History of		Periodic
	τ	m	Revelation
Summary Statistics.			
Consumer Surplus	2.05 e+07	2.05 e+07	2.05 e+07
Profit B	2.46 e+06	2.46 e+06	2.45 e+06
Profit S	2.32 e+06	2.32 e+06	2.33 e+06
Maintenance Cost B	2.28 e+05	2.28 e+05	2.28 e+05
Maintenance Cost S	1.66 e+05	1.66 e+05	1.65 e+05
Production Cost B	2.40 e+06	2.40 e+06	2.39 e+06
Production Cost S	2.82 e+06	2.83 e+06	2.83 e+06

The Social Planner Problem. The solution to the social planner problem provides a basis for understanding the logic underlying efficient maintenance decisions for our parameterization. Recall that there is significantly less demand on weekends than on weekdays. Table 4 presentsaverage shutdown probabilities by day of week. The social planner shuts down at least one large and one small generator about 97% of the Sundays, and shuts down two of each type of generator over 60% of all Sunday's. As a result Monday is the day with the maximum average number of both small and large generators operating. The number of generators operating falls on Tuesday, and then falls again both on Wednesday and on Thursday, as the cost state of the generators maintained on Sunday stochastically decay and maintenance becomes more desirable. By Friday the planner tends to favor delaying further maintenance until the weekend, so the number of generators operating rises. Maintenance goes up slightly on Saturday, but there is an obvious planner preference for doing weekend maintenance on Sunday, as that enables the generators to be as prepared as possible for the Monday work week. As Table 5 shows these maintenance decisions imply that almost no maintenance occurs at low cost states ($\omega = 0$ or $\omega = 1$).

The Monopoly Problem. The solution to the monopoly problem provides a basis for understanding the logic of profit maximizing behavior when the two firms can fully co-ordinate their behavior. Profit maximization gen-

Table 4: **Average No. of Operating Generators.**

	Weekend		Weekdays				
	Sat.	Sun.	Mon.	Tues.	Wed.	Thurs.	Fri.
Social Planner							
Firm B:	2.38	1.24	2.85	2.44	2.08	2.06	2.43
Firm S:	2.79	2.08	3.11	3.08	2.96	2.96	3.12
Duopoly AI							
Firm B:	2.17	2.16	2.29	2.32	2.24	2.22	2.27
Firm S:	3.32	2.91	2.16	2.41	2.57	2.50	2.50
Duopoly FI							
Firm B:	2.02	1.81	1.80	1.84	1.87	1.81	1.84
Firm S:	2.62	2.43	2.35	2.42	2.41	2.40	2.42
Monopolist							
Firm B:	3	2.62	0.97	1	1	1	1
Firm S:	3.90	3.39	0.03	0	0	0	0

erates a very different outcome than does the surplus maximization of the planner. On virtually every weekday the monopolist shuts down two of the three large generators and *all* of the small generators. In contrast, on weekends it tends to operate all three of its large generators, and at least three of the four small generators. That is the monopolist has an incentive to hold back on supply when demand is high, and supply large amounts when prices are low. Moreover this generates a shut down pattern that seems socially inefficient; 43% of the shutdowns of large generators and over 80% of the shutdowns of small generators occur at the most favorable cost state of $\omega = 0$. Indeed the cost state of the monopolist's operating generators is, on average, half or less that of the social planner's operating generrators.

The Duopoly with FI. When there is full information the aveage number of generators operating is close to constant over the whole week (weekday or weekend; though on Saturday utilization rates do increase a small amount for both firms). Indeed the full information AME solution leaves the two firms with one of two combinations of operating generators over 70% of the time on *each* weekday; about 45% of the time there are two of each type

Table 5: **Distribution of ω Prior to Shutdown.**

	Dist. ω Prior to Shutdown.					**Maint***
	$\omega = 0$	$\omega = 1$	$\omega = 2$	$\omega = 3$	$\omega = 4$	**Freq.**
Social Planner						
Firm B:	0.00	0.002	0.070	0.152	0.778	2.81
Firm S:	0.00	0.012	0.150	0.250	0.588	2.60
Duopoly AI						
Firm B:	0.021	0.010	0.020	0.026	0.924	2.94
Firm S:	0.201	0.076	0.150	0.122	0.452	1.97
Duopoly FI						
Firm B:	0.182	0.158	0.267	0.080	0.313	1.62
Firm S:	0.270	0.120	0.181	0.102	0.327	1.55
Monopolist						
Firm B:	0.432	0.135	0.216	0.056	0.162	.97
Firm S:	0.811	0.026	0.061	0.064	0.040	.36

* Average number of days between maintenance decisions.

of generator operating, and about 26% of the time three of each type of generator is operating.

This leads to less shutdown occuring at low costs states than we found for the monopolist. However still well over a third of the shutdown decisions for each type of generator occur when the generator is at one of the two lowest costs states (the planner has almost no shutdowns at those states). Moreover the full information duopoly firms do maintenance about 70% more than does the social planner, and supply a bit more electricity on weekends then on weekdays. We come back to the welfare implications of this behavior below.

The Duopoly with Asymmetric Information. Perhaps the most striking result is in Table 5 is the finding that the frequency of maintenance in the equilibrium of the AI duopoly is similar to that frequency in the social planner solution. The firm with the big generators in the AI duopoly does a little less maintenance than a social planner would, and the firm with the small generators does a little more (see table 5). This contrasts sharply with the behavior of the firms in the equilibrium of the FI duopoly.

Relatedly the distributions of ω prior to shutdown are almost stochastically orderred (in the first order dominance sense) across institutional regimes (see table 5). The distribution of maintenance emanating from the AI duopoly is pretty close to being larger than that of the planner, which *is* larger than that generated by the FI duopoly, which in turn is larger than the monopoly generates. So to the extent there is an inefficiency in the maintenance decisions of the asymmetric information equilibrium it *does not* seem to be a result of firms *withdrawing* too much capacity; if anything the firm with the large generators in the duopoly with asymmetric information *does not do enough* maintenance.

This is strikingly different from the maintenance decisions in the FI AME equilibrium, wherein there is clearly too much maintenance by both firms. There are two possible reasons for this difference; part of the difference in maintenance behavior could be a result of the differences in AME recurrent class generated by the two equilibria, and part could be a result of differences in strategies for a given set of states. To get some indication of the relevance of these two possiblities we took the invariant distribution of states in the two AME equilibria and considered the bids that would be generated by the *static* Nash full information equilibrium in those states. When using the static Nash full information policies the number of generators operated did not differ across the states generated by the two equilibria.[24] Apparently the uncertainty generated by asymmetric information induces firms to do less maintenance.

We come back to the consumer welfare implications of this presently. To see why the firms bid more generators (do less maintenance) in the AI than in the FI equilibrium, consider the incremental profits on the next machine firm j could have bid in the FI equilibrium. That is consider

$$\Delta\pi(m_j^o + 1, m_{-j}^o, d_j, \omega_j) \equiv \pi_j(m_j^o + 1, m_{-j}^o, d_j, \omega_j) - \pi_j(m_j^o, m_{-j}^o, d_j, \omega_j),$$

where m_j^o is the number of machines bid. Now consider the incremental profits of player j at the same (m_j^o, ω_j) in the AI model. In the AI model

[24]The average number of small generators operating in a day was 2.85 from both sets of states and for the large generator was 2.22 for the distribution of states from FI dynamic duopoly versus 2.27 for the distribution of states from the AI dynamic duopoly. When there were multiple equilbria to the static game these numbers were computed by averaging over those equilibria.

m^o_{-j} is unknown, so the decision of firm j is based on an expectation over its likely values. The variance introduced by the asymmetric information will increase the profitability of the $m^o_j + 1$ generator if $\Delta\pi(\cdot, m_{-j}, \cdot)$ is convex in m_{-j}. That is the effect of increased variance on the quantity bid in a static Nash equilibria will depend on the convexity of the first order condition in the quantities bid by a firm's competitors.

In a dynamic game maintenance decsions are based on the value function (not on the profit function). So whether or not the introduction of asymmetric information increases the incentives to bid generators depends on the convexity of the value function in the number of generators bid by the firm's competitors. However the value function is a complex iterate of the profit function and we might expect it to inherit the properties of that function. To check whether this was so in our case we calculated the increment in a firms' profit for bidding an additional generator for all points which were in the recurrent class of both the FI and AI AME equilibria. We then looked to see whether this increment was the discrete analogue of "convex" in m_{-i}. It was, and distincly so (at least for the averages we printed out). The extent to which this result generalizes to other specificaitons is of more general interest to the analysis of Nash equilibria with asymmetric information.

Though the extent of shutdown in the AI AME equilibria is similar to the extent to which the social planner shuts down generators, as table 2 shows, when there is asymmetric information shutdowns occur more on weekends than on weekdays, just the opposite of the social planner. That is the duopoly with asymmetric informations' distribution of maintenance decisions over days of the week is not efficient. Moreover the equilibrium with asymmetric information sometimes incentivizes the firm with the small generators to shut down the "wrong" generators; i.e. to shut down generators with lower cost states than generators that it operates (something which never occurs in any of the other regimes, and almost never occurs to the firm with the large generators in the duopoly with AI).

We considered all those cases where the firms shut down 1 or 2 generators and computed the fraction of those times that it shut down the generators with the highest cost state (the highest value of ω). In the duopoly with AI, when the firm with the small generators shut down one generator it *did not* shut down the highest cost generator 30% of the time, and when two generators were shut down it did not shut down the two highest cost generators over 35% of the time. Since, if the competitor did not change its bid, the firm would always do better by shutting down the highest cost generator, this can only occur because by shutting down a higher cost generator the firm induces its competitor to change its bids in the coming period in a way that favors the firm which is shutting down the low cost generator; i.e. because of the signalling value of the maintenance decision. Interestingly when the firm did not shut down the generator with the highest cost the low cost generator that was shut down had been operating less time since its last maintenance decision then the high cost generator that was not. So by shutting down the low cost generator the firm insures that the generators that it will operate next period will have operated a longer period of time since their last maintenance decision.

Table 6: **Welfare Under Alternative Institutions.**

	Duopoly AI	Duopoly FI	Planner	Monopolist
Cons. Surplus (CS) ($\times 10^{-6}$)	20.51	19.70	22.21	2.05
Profits.				
Firm B ($\times 10^{-6}$)	2.45	2.11	1.99	11.52
Firm S ($\times 10^{-6}$)	2.33	2.83	2.13	-.250
Firms B + S ($\times 10^{-6}$)	4.78	4.95	4.12	11.27
Total Surplus (CS + B + S)	25.29	24.65	26.34	13.32
Prices.				
Weekend	145.77	170.42	152.51	216.44
Weekday	1205.76	1292.83	990.46	3880.02
Fraction of Output Produce by Firm with Larger Generators.				
Weekend	.47	.48	.46	.60
Weekday	.50	.43	.46	.99

Consumer Surplus and Profits. As one might expect there is a rather large difference in both consumer suplus and profits between the monopolist and the social planner (see Table 6). The fact that the monopolist holds back generators so much of the time implies that consumer surplus is about 10.8 times higher under the social planner's allocation (22.21 vs 2.05). Some of this difference gets made up in profits; the discounted profits under the monopoly allocation are about two and half times those under the social planner's allocation. However the difference in profits does not come close to compensating for the consumer surplus loss; total surplus is about twice as large under the social planner's allocation (26.34 vs 13.32).

This large a difference in total surpls between the monopoly and social planner's allocations implies that these primitives leave a lot of room for differences in the performance of different institutional structures. However both our duopoly equilibria, that with asymmetric and that with full information, generate a total surplus which is much closer to that from the social planner problem. The full information AME equilibrium generates a total suplus just 6.5% less than that of the social planner, while the asymmetric information AME equilibrium does even better; generating a surplus which is only 4% less than the social planner does. As one might expect from our discussion of maintenance decisions consumer surplus is larger under AI AME equilibria than under the FI AME equilibria, but the reverse is true for firm profits.

Note that even in the solution to the social planner problem we see rather dramatic price effects of the differential demand between weekdays and weekends. Both the FI and the AI AME equilibria magnify this difference, but not to anywhere near the extent of the monopoly equilibria. Note also that it is the firm with the small generators' who gains the most from moving to full information. This is because when there is full information the firm with small generators produces a higher fraction of the output on the highly lucrative weekdays (58% vs 50%), and has costs which fall more than proportionately to its decrease in production (because without asymmetric information it always produces with its most efficient generators). That is by moving to full information the firm with relatively low maintenance costs but high production costs is able captures more of the surplus available on high demand days.

5 Concluding Remark

We have presented a simple framework for analyzing finite state dynamic games with asymmetric information. It consists of a set of equilbrium conditions which, at least in principal, are empirically testable, and an algorithm capable of computing equilibrium policies from a given set of primitives. 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. There are many dynamic situations of interest to Industrial Organization which naturally involve asymmetric information; examples include collusion, auctions, and regulation. A more in depth analysis of these situations requires a framework which is relatively easy to use and can incorporate the empirical detail that insures a realistic description of the phenomena of interst.

References

Bajari, Benkard and Levin, 2007, “Estimating Dynamic Models of Imperfect Competition”, *Econometrica*, September.

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

Borenstein S., J. Bushnell, and F. Wolak (2002) ” Measuring Market Inefficiencies in Californaia’s Restructured Wholesale Electricity Market, *American Economic Review*, pp. 1376-1405.

Dekel, E., D. Fudenberg and D.K. Levine (2004), ”Learning to play Bayesian Games” *Games and Economic Behavior*, 46, 282-303.

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

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

Fudenberg, D and D. Kreps (1994) ”Learning in Extensive Form Games. II Experimentation and Nash Equilibrium” Mimeo Stanford University.

Fudenberg, D. and D.K. Levine (1993a) ”Self Confirming Equilibrium” *Econometrica*, 61(3), 523-545.

Fudenberg, D. and D.K. Levine (1993b) ”Steady State Learning and Nash Equilibrium” *Econometrica*, 61(3), 547-573.

Harvey, S, Hogan, W., and Schatzki, T. (2004) ” A Hazard Rate Analysis of Mirant’s Generating Plant Outgaes in California” *mimeo*, the Kennedy School of Government.

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

Maskin, E. and J. Tirole (1988), “A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs” *Econometrica*, 56, 549-570.

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.

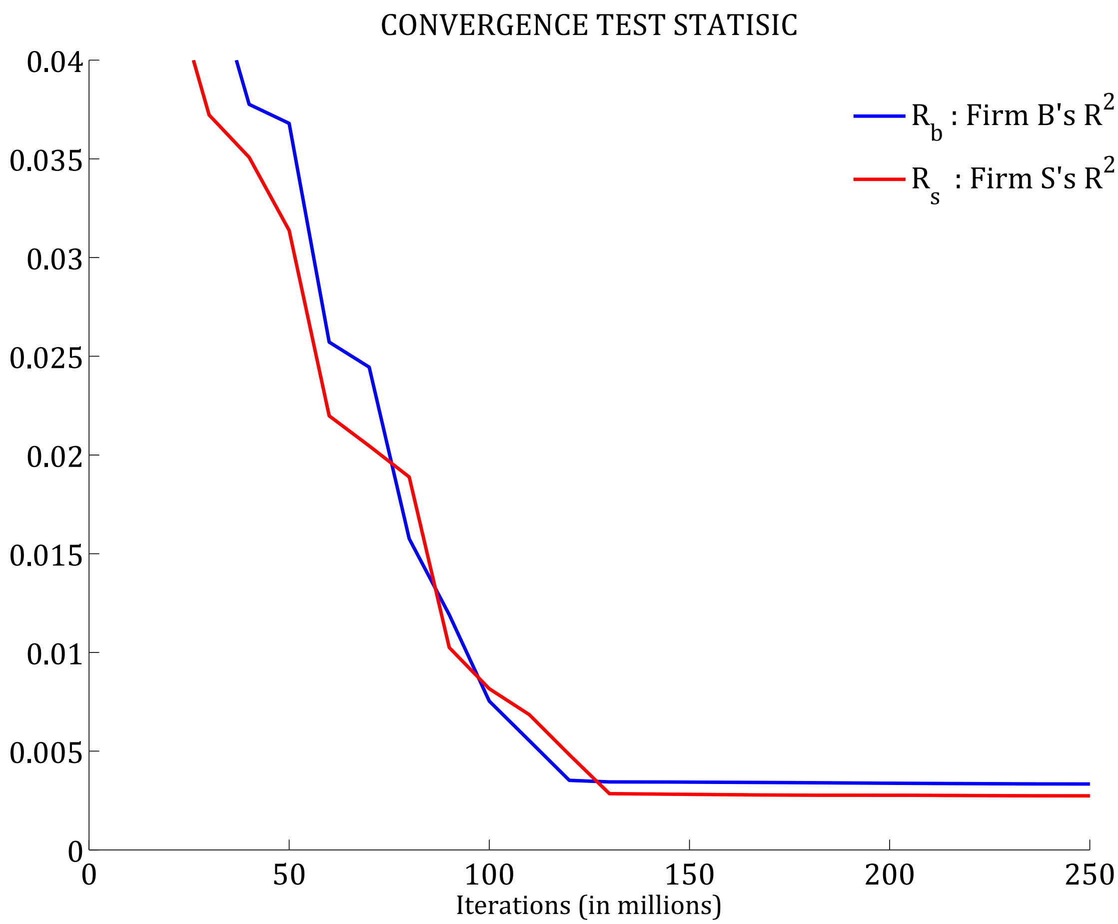
Pakes, A. and P. McGuire (1994), "Computing Markov Perfect Nash Equilibrium: Numerical Implication of a Dynamic Differentiated Product Model", *The Rand Journal of Economics*, 25, pp. 555-589.

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

Pakes Ostrovsky and Berry (2007), "Simple Estimators for the Parameters of Discrete Dynamic Games with Entry Exit Examples", *RAND Journal of Economics* v.38(2), Summer 2007, pp.373-399.

Robbins and Monroe (1954); "A Stochastic Approximation Technique", *Annals of Mathematics and Statistics*,.

Wolak, F., and R. Patrick (1997); "The Impact of Market Rules and Market Structure on the Price Determination Process in the Engaland and Wales Electricity Market" *POWER* working paper PWP-047, Unicersity of California Energy Institute ", *Annals of Mathematics and Statistics*,.


CONVERGENCE TEST STATISIC
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
50
100
150
200
250
Iterations (in millions)
R_b : Firm B's R^2
R_s : Firm S's R^2