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

Chaim Fershtman and Ariel Pakes
(Tel Aviv University and Harvard University).

May 26, 2009

Abstract

With applied work in mind, we define an equilibrium notion for dynamic games with assymmetric 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.

*We would like to thank John Asker, Susan Athey, Adam Brandenburger, Eddie Dekel, Liran Einav, Drew Fudenberg, Phil Haile, Ehud Kalai, Naryana Kocherlakota, Robin Lee, Greg Lewis and Michael Ostrovsky, for helpful comments, and Niyati Ahuja for superb research assistance.

This paper develops a relatively simple framework for the analysis of dynamic games with sources of asymmetric information whose impacts persist over time. We consider a class of dynamic games in which there are a finite number of players in each period (the number may change over time with entry and exit). Each firm's profits in a given period are determined by all firms' "payoff relevant" state variables and their actions. Neither a player's "payoff relevant" state variables (e.g. indexes of their cost function, qualities of the goods they market, etc.) nor its actions are necessarily observable to the other firms. Accordingly, in addition to payoff relevant state variables firms have "informationally relevant" state variables which provide information on the likely play and/or states of its competitors. The state variables of each firm evolve over time with the actions of all firms and the random realizations of exogenous processes. So each player's action may affect any players; profits, payoff relevant state variables, and/or infromationally relevant state variables (as would occur if firms sent signals.).[1]

Our goal is to develop a framework for the analysis of dynamic games with persistent sources of assymetric information which can be used in applied work. Consequently 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 we show that an equilibrium that is consistent with a given set of primitives can be computed using a simple reinforcement learning algorithm. Neither the iterative procedure which defines the algorithm nor the test of the equilibrium conditions are subject to a curse of dimensionality[2].

[1]Dynamic games with asymmetric information have not been used extensively to date, a fact which attests (at least in part) to their complexity. Notable exceptions are Athey and Bagwell, 2008, and Cole and Kocherlakota (2001). Athey and Bagwell consider collusive Markov perfect equilibria when the cost position of each firm is not observed and evolves as an exogenous Markov process. Cole and Kochelarkota focus on "Markov-private" equilibria in which equilibrium strategies are constrained to depend on private information only through the privately observed state.

[2]Reinforcement learning 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, a fact which has been used in several applied papers; e.g. Goettler, Parlour, and Rajan, 2005, and Beresteanu and El-

One could view our reinforcement learning algorithm as a description of how players' learn the implications of their actions in a changing environment. This provides an alternative reason for interest in the output of the algorithm. However it also accentuates the fact that our framework is only likely to provide an adequate approximation to the evolution of a game in which it is reasonable to assume that agent's perceptions of the likely returns to their actions can be learned from the outcomes of previous play. Since the states of the game evolve over time and the possible outcomes from each action differ by state, if agents are to learn to evaluate these outcomes from prior play the game needs to be confined to a finite space. We consider different ways of insuring finiteness, the simplest of which is to truncate the history an agent is able to keep in its information set.

We define a state of the game to be the information sets of all of the players (each information set contains both public and private information). An Applied Markov Equilibrium (hereafter, AME) for our game is 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 each firm's expected discounted value of net cash flow conditional on the possible actions that agent can take. The conditions we impose on this triple are as follows. 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 by 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 of outcomes used to form strategies be consistent with competitors' play when that play results in outcomes that are in the recurrent class of points, and hence are observed repeatedly. We allow for feasible outcomes that are not in the recurrent class, but the conditions we place on the evaluations of those outcomes are weaker; they need only satsify inequalities which insure that they are not observed repeatedly. In this sense

lickson, 2006. Goettler, Parlour, and Rajan, 2008, use it to approximate optimal behavior in a trading game with asymmetric information.

our notion of equilibrium can be considered an extension of the notion of Self Confirming equilibrium, as defined by Fudenberg and Levine (1993a)[3], to dynamic games; a point we come back to below. An implication of using the weaker equilibrium conditions is that we might admit more equilibria than the Markov Perfect concept would.

To illustrate we conclude with an example that endogenizes the maintenance decisions of electricity generators. We take an admittedly simplified set of primitives and compute and compare equilibria based on alternative institutional constraints. These include; asymmetric information equilibria where there are no bounds on agents memory, asymmetric information equilibria where there are such bounds, symmetric information equilibria, and solutions to the social planner and monopoly problem. We use the results to consider whether the bounded memory constraints cause differences in behavior and to compare asymmetric to symmetric informational environments. The results also provide some insights into the relationship between strategic withdrawl of generators and electric utility prices.

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 and test for 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.[4] 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 determinants 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 of all the firms,

[3]See Kalai and Lehrer, 1993, for a similar equilibrium concept.

[4]Indeed our assumptions nest the generalizations to Ericson and Pakes (1995) reviewed in Doraszelski and Pakes(2008). For more detail on existence and uniqueness of equilibria in these games, see Doraszelski and Satterthwaite, 2007.

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; 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 other firm's profits (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 actions can affect current profits and/or the probability distribution of payoff relevant random variables[5].

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 (these account for demand and input price movements over time). 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

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

[5]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.

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

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

is partially controlled by firm's choice of m and x and can be influenced by the m choices of other firms. The distribution of ν is given exogenously and equal to

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

Capital accumulation games, that is games where the evolution of the firm's own state variable depends only on the firm's own actions, are a special case of equation (3) in which $m \in M$ or

$$\mathcal{P}_{\eta} = \{ \ P_{\eta}(.| \ m, x, \omega); \ \ m \in \mathcal{M}, \ \ x \in X, \ \ \omega \in \Omega \}. \tag{4}$$

Both the equilibrium conditions and the algorithm for computing equlibria are noticeably simpler for the specification in equation (4) then for the general case in (3). On the other hand the generality in (3) is needed for many cases of interest. One example is a learning by doing games in which ω represents marginal cost which is private information and evolves over time as a function of the firm sales, m refers to prices, and firms control prices. Then the choice of a firm's m affects the evolution of *all* firms' costs. As discussed below many dynamic auction situations also require the generality in (3) as then each firm's bid (our m) effects the probability of any firm winning the current auction and winning the current auction often effects the state of the firm going into the next auction. To facilitate understanding while at the same time providing the needed generality we deal with the special case of equation (4) in the text, but provide the added detail needed to analyze the more general case in (3) in the appendix[6].

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

[6]To accomodate the generality in (3) we need to to modify both the notion of equilibrium and the iterative algorithm given in the body of the paper. This generality, however, allows us to nest the asymetric information analogues of all the modifications of the Ericson and Pakes (1995) model reviewed in Doraszelski and Pakes (2008).

follows. The "payoff relevant" variables are defined, as in Maskin and Tirole (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, or (ii) even if player i's strategy does not condition on the variable there is at least one player j whose 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.[7] 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 of the agent's information set, so we will need separate notation for them. The publicly observed component will be denoted by $\xi_t \in \Omega(\xi)$, while the privately observed component will be $z_{i,t} \in \Omega(z)$. For example $\omega_{-i,t-1}$ may or may not be known to agent i at time t; if it is known $\omega_{-i,t-1} \in \xi_t$, otherwise $\omega_{-i,t-1} \in z_{-i,t}$. 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 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. Formally if ϵ_{t+1} denotes the public information available in period $t+1$ but not in t, we assume that

$$\xi_{t+1} = G_\xi(\xi_t, \nu_{t+1}, \epsilon_{t+1}), \tag{5}$$

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

$$\mathcal{P}_\epsilon = \{ \; P_\epsilon(\cdot | \; \xi, x, m, z, \eta), \; (\xi, x, m, \eta, z) \in \Omega(\xi) \times (X \times \mathcal{M} \times \Omega(\eta) \times \Omega(z))^n \; \}. \tag{6}$$

[7]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.

These transition probabilities will depend both on the stochastic specification given above, and on the institutional structure that determines how information is revealed. For example if $\omega_{i,t} \in z_{i,t}$ but *is revealed* by agent i's play in period t, then $\omega_{i,t} \in \epsilon_{t+1}$, and the distribution of ϵ_{t+1} depends on $z_{i,t}$. We shall assume that $G_\xi(\cdot)$ deletes public information which was either payoff or informationally relevant in t but no longer is in $t+1$.

Similarly if $\mu_{i,t+1}$ denotes the private information available in period $t+1$ but not in t we assume that

$$z_{i,t+1} = G_z(\xi_t, z_{i,t}, \mu_{i,t+1}), \tag{7}$$

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

$$\mathcal{P}_\mu = \{ \ P_\mu(\cdot| \ \xi, x, m, z, \eta), \ (\xi, x, m, \eta, z) \in \Omega(\xi) \times X \times \mathcal{M} \times \Omega(\eta) \times \Omega(z) \ \}. \tag{8}$$

For example if $\eta_{i,t+1}$, the stochastic increment in $\omega_{i,t+1}$, is only seen by firm i then $\eta_{i,t+1} \in \mu_{i,t+1}$. We also assume that $G_z(\cdot)$ 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.

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

For concreteness we assume timing of the game is as follows (this timing can be changed with appropriate modifications to the equilibrium conditions and the computational algorithm introduced below). At the beginning of each period there is a realization of $\{\mu, \nu, \epsilon\}$. Firms then update their information sets with the updating functions (7) and (5). 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) symmetric information model is the special case where ξ_t (the public information) is $(\omega_{i,t}, \omega_{-i,t})$ and its increment is $\epsilon_t = (\eta_{i,t}, \eta_{-i,t})$. A simple asymmetric information model which has been used extensively in recent empirical work is one in which the public information is the same as above and, in addition, each firm has a cost shock which is private information and is independently distributed over time, so $z_{i,t} = \mu_{i,t}$ (see Bajari, Benkard and Levin, 2007, and Pakes, Ostrovsky and

Berry, 2007). The electric utility model with endogenous maintenance decisions for generators analyzed in section (4) of this paper allows for asymmetric information on a cost component which is serially correlated over time, bids which contain signals on those costs, and actions which make those costs public information.

Some sequential auctions fit into the capital accumulation framework of assumption $\mathcal{P}_\eta$ in equation (4), while others require the generality of assumption $\mathcal{PG}_\eta$ in equation (3). For an example which abides by $\mathcal{P}_\eta$ consider a sequence of procurement auctions for high technology products (e.g. medical instruments or major weapons systems). Let ω represent the level of technology of the bidder which is private information and evolves with the outcome of the firm's R&D activity (our x). The bids (our m) determine who wins the contract while at the same time sends a signal on the bidder's ω. Since the bid of one firm does not effect the transitions of other firm's technology states the capital accumulation assumption in $\mathcal{P}_\eta$ is relevant.

On the other hand consider a sequence of timber auctions. Each competitor has a fixed capacity of processing harvested timber (which is public information), and each period there is a new lot up for auction. The quantity of timber on the lot auctionned is unknown at the time of the auction and only revealed to the firm that submits the highest bid and wins the auction. The amount of unharvested timber on the lots the firm owns is private information (our ω). Each period each firm decides how much to bid on the current auction (our m) and how much of its unharvested capacity to harvest (our x). The timber that is harvested and processed is sold on an international market which has a price which evolves exogenously (our $\{d_t\}$ process). Let l_t be the quantity of timber on the lot auctioned at time t, and $\eta_{i,t} = \{m_{i,t} = max_j m_{j,t}\} l_t$ where $\{\cdot\}$ is an indicator function that takes the value of one if the condition inside the curly brackets is satisfied and zero elsewhere. Then $\omega_{i,t+1} = \omega_{i,t} - x_{i,t} + \eta_{i,t}$, and to formulate the η distribution we require the generality in assumption $\mathcal{PG}_\eta$. The fact that many applied examples require this generality is the reason our appendix provides an extension that enables them to be analyzed.

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 $\mathcal{S} = \{s : z \in \Omega(z)^n, \xi \in \Omega(\xi),$ 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 distribution of 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. 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 the union of the information sets available to the agents playing the game. We come back to these issues, and their relationship to past 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 $\eta(x^*(J_i), m^*(J_i), \omega_i) \equiv \{\eta : p(\eta|x^*(J_i), m^*(J_i), \omega_i) > 0\}$; the set of η values that have positive probability when equilibrium strategies are played. For every $\eta \in \eta(x^*(J_i), m^*(J_i), \omega_i)$

$$W(\eta, m^*(J_i)|J_i) = \pi^E(J_i) + \beta \sum_{J_i'} \left\{ \sum_{\tilde{\eta}} W(\tilde{\eta}, m^*(J_i')|J_i') p(\tilde{\eta}|x^*(J_i'), m^*(J_i'), \omega_i') \right\} p^e(J_i'|J_i, \eta),$$

where

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

and

$$\left\{ p^e(J_i'|J_i, \eta) \equiv \frac{p^e(J_i', \eta|J_i)}{p^e(\eta|J_i)} \right\}_{J_i'}, \quad \text{and} \quad \left\{ p^e(J_{-i}|J_i) \equiv \frac{p^e(J_{-i}, J_i)}{p^e(J_i)} \right\}_{J_{-i}} \spadesuit. \quad (9)$$

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. Similarly 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. That is the $W(\eta, m|J_i)$ of an η_i not a component of an $\eta \in \eta(x^*(J_i), m^* J_i, \omega(J_i))$ or of an $m \neq m^*$ 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.[8]

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.

Comment 1. Applied Markov Equilibria, as defined above, can be thought of as an extension of Self Confirming Equilibria (Fudenberg and Levine,1993a) to dynamic games.[9] Self Confirming Equilibria weaken the standard Nash equilibrium conditions. They require that each player has beliefs about opponents' actions and that the player's actions are best responses to those beliefs. However the players' beliefs need only be correct along the equilibrium path. This insures that no players observes actions which contradicts its beliefs. Our equilibrium conditions explicitly introduce the evaluations that the agents use to determine the optimality of their actions. They are similar

[8]The fact that our conditions do not apply to points outside of $\mathcal{R}$ or to $\eta \notin \eta(x^*(J_i), m^*(J_i), \omega(J_i))$ implies that the conditional probabilities in equation (5) are well defined.

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

to the conditions of Self Confirming Equilibria in that the most they insure is that these evaluations are consistent with the opponents actions along the equilibrium path. However we distinguish between states that are repeated infinitely often and those that are not, and we do not require the evaluations which determine actions at transitory states to be consistent with the play of a firm's opponents.

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 $W(\eta, m|J_i)$ for all $m \in \mathcal{M}, \eta \in \Omega(\eta)$, and J_i which is a component of s. Then condition C2 tests for the optimality of strategies when all possible outcomes are evaluated in a way which is insured to be consistent with all the players' equilibrium actions. At an s which is a boundary point we only obtain consistent estimates of the $\{W(\cdot)\}$ for $(\eta, m) \in ((\eta(x^*(J_i), m^*(J_i), \omega(J_i)), m^*(J_i)))$ for each J_i component of s, that is for the states which are observed with positive probability and equilibrium actions. Then condition C2 tests for optimality in the restricted sense that only some of the outcomes are evaluated in a way that is insured to be consistent with competitors' equilibrium play (for more detail see section 3.2 below).[10]

[10] Two additional points are worth noting here. First if, at a boundary point, an agent can take an action which can lead to a point outside of the recurrent class at which there is a known minimum value regardless of the actions or states of competitors (e.g. the value if it exited), then one could add the condition that the evaluation of the $W(\eta, m|J)$ for that (η, m, J) combination must be at least as high as the minimum value. Second the empiricist may well observe less than s_t, perhaps only the publically available information in each period. Provided the empiricist knows (or has estimates of) the primitive parameters, testing would then consist of asking whether the observed probabilities of transition from one public information set to another are consistent with an equilibrium set of policies.

Comment 3. As noted there may be more than one AME associated with a given set of primitives and the fact that we have not considered any restrictions on the evaluations outside of the recurrent class may contribute to the possible mulitplicity. Section 3 of this paper considers ways of directing the computational algorithm towards particular equilibria. Note, however, that there are (generically) unique equilibrium strategies associated with any given equilibrium $\{W(\cdot)\}$, so if we can assume equilibrium play and estimate the $\{W(\cdot)\}$ associated with the set of $s \in \mathcal{R}$ we can compute the entire distribution of sample paths from any given point in $\mathcal{R}$.

2.1 The Finite State Space Condition.

In our definition of an AME we restricted ourselves to games where the state space was finite, and the next section provides an algoirthm which computes AME's for games with finite state spaces. We now consider this conditions in more detail.

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[11]. In our context 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 history of these variables.

There are a number of ways to satisfy our finite state space condition. One is to consider a class of games in which the agents would not gain by dis-

[11]As in Ericson and Pakes (1995). These conditions 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 independently distributed private information took on an infinite range of values, then the value functions could take on an infinite range of values, but given the rest of our assumptions the continuation values would lie in a finite set.

tinguishing 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 in Claim 1 below.

A second way to insure finiteness is to assume there are bounded cognitive abilities and these bounds generate a finite state space. This may be an appealing restriction to applied researchers since in most applications the available data is truncated and various constraints limit players ability to condition their strategies on large state spaces. One obvious example of such an assumption would be a direct bound on memory (for e.g. 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.[12]

A third way to insure finiteness is to consider institutional structures wherein each agent only has access to a finite history. For example consider a sequence of internet auctions, say one every period, for different units of a particular product. Potential bidders enter the auction site randomly and can only bid at finite increments. Their valuation of the object is private information, and the only additional information they observe are the sequence of prices that the product sold at while the bidder was on-line. If, with probability one, no bidder remains on the site for more than L auctions, prices more than L auctions in the past are not in any bidder's information set, and hence can not effect bids.[13]

In our computational example we compute equilibria to finite state games generated by the first two types of assumptions. Indeed one of the questions we adress is whether the different assumptions we use to obtain finiteness, all of which seem *a priori* reasonable, generate equilibria with noticeably different policies. As we shall see for our example they do not.

The example of institutions which insure that the equilibrium we compute for a finite state space is an equilibrium to the unrestricted game used in our computational example is the case of periodic simultaneous revelation of all

[12]Note that any empirical application that used a bound on memory could examine empirically whether relaxing the bound has any noticeable effect on policies.

[13]Formally this example requires an extension of our framework to allows for state variables that are known to two or more, but not to all, agents. Further to insure that our AME equilibrium provided a reasonable approximation to behavior in this market we would want to point to a mechanism which enabled bidders to form equilibrium perceptions of winning conditional on bids and observed prices.

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 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 valuations, strategies, and probability distributions when agents condition both their play and their expectations (which are their valuations) 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 (7) and (6) 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.[14] 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[15].

[14] This is similar to 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.

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

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[16]. We conclude this section with a brief discussion of the properties of the algorithm.

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 literature (a literature which dates to Robbins and Monroe, 1954). The algorithm is iterative. An iteration, say k, is defined by a couple consisting of

- its location, say $L^k = (J_1^k, \ldots J_{n(k)}^k) \in \mathcal{S}$, which is defined by the information set of the $n(k)$ agents active at iteration k[17], 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 (4), (6) and (8)) conditional on those policies and $L^k = (J_1^k, \ldots J_{n(k)}^k)$ are taken. Those draws are used to update both L^k and W^k.

[16]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 experience in related states.

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

The location is updated using the updating functions in equations (5) (for the public information) and (7) (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 a weighted average of: (i) the values at the updated state (as determined by the evaluations in memory) and (ii) the initial values at L^k (i.e. the W^k). The update generates a W^{k+1} which is a weighted average of the values obtained from all iterations which started at the same location as did iteration k. We now formalize this procedure and then discuss some of its properties.

Details. The stochastic approximation 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 $M^k = \{M^k(J) : J \in \mathcal{J}\}$.

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 (if d is random, then the algorithm has to take a random draw on it before calculating profits). 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 4, 6 and 8), 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 (5) and (7)) to obtain the updated location of the algorithm

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

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

$$V^{k+1}(J_i^k) = \pi(\omega_i^k, \omega_{-i}^k, m_i^k, m_{-i}^k, x_i^k, d^k) \ + \qquad (10)$$

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

Note that to calculate $V^{k+1}(J_i^k)$ we need to first find and call up the information in memory at locations $\{J_i^{k+1}\}_{i=1}^{n_{k+1}}$.[18] Once these locations are found we keep a pointer to them, as we will 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 (10) 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

[18]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.

$$W^{k+1}(\eta, m|J_i^k) - W^k(\eta, m|J_i^k) = \frac{1}{A\Big(h^k(J_i^k)\Big)}[V^{k+1}(J_i^k) - W^k(\eta, m|J_i^k)], \quad (11)$$

where $A(\cdot) : \mathcal{Z}^+ \to Z^+$, is increasing, and satisfies Robbins and Monroe's conditions (1951)[19]. 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 above to find the information stored in memory at L^{k+1}. This completes the iteration as 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 determine 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 $q(s|\tilde{W}, \tilde{p})$. 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

[19] 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.

kernel and so any sample path generated by it and any initial condition will enter a recurrent class with probability one in a finite number of periods, and once within that class the process will stay within it forever. Let $\tilde{\mathcal{R}} \subset \mathcal{S}$ be one of the possible recurrent classes. Our test consists of generating an $\tilde{\mathcal{R}} \subset \mathcal{S}$ and testing whether the $\{\tilde{W}(s);\ s \in \tilde{\mathcal{R}}\}$ satisfy the equilibrium conditions (C2 and C3 above).

To obtain our candidate for $\tilde{\mathcal{R}}$, 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. The recurrent subset of points obtained in this way is our candidate for $\tilde{\mathcal{R}}$. 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.[20]

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 $s \in \tilde{\mathcal{R}}$. 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 needed probability distributions, $p^e(J_i'|J_i)$ and $p^e(J_{-i}|J_i)$. 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. They show that 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.

[20]An alternative, but computationally more burdensome, way of finding a recurrent class generated by $\tilde{W}$ would be to compute all transitions implied by the optimal policies and test for a recurrent class associated with each (or a subset of the) $s \in \mathcal{S}$. Though computationally more burdensome, this procedure should allow the researcher to find all possible recurrent classes generated by $\tilde{W}$.

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 computing the differences between the estimates of $\tilde{W}(\eta, m^*(\tilde{W})|J_i)$ in memory, and a simulated 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 (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 $\tilde{W}$ (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 a 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 $A\hat{S}W^1(\cdot)$, 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 that condition on a particular initial location, then

$$E\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)} - 1\Big)^2 = E\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i) - E\Big[A\hat{W}(\eta_i, m^*(J_i)|J_i)\Big]}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big)^2 + \Big(\frac{E\Big[A\hat{W}(\eta_i, m^*(J_i)|J_i)\Big] - \tilde{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big)^2. \tag{12}$$

The first term after the equality in (12) 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

$$V\hat{a}r\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big) \equiv \frac{A\hat{S}W(\eta_i, m^*(J_i)|J_i)\frac{h(\eta_i,J_i)}{h(\eta_i,J_i)-1} - A\hat{W}(\eta_i, m^*(J_i)|J_i)^2}{\tilde{W}(\eta_i, m^*(J_i)|J_i)^2}, \tag{13}$$

then

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

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

$$B\hat{i}as(AW(\eta_i, m^*(J_i)|J_i))^2 \equiv \Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)} - 1\Big)^2 - V\hat{a}r\Big(\frac{A\hat{W}(\eta_i, m^*(J_i)|J_i)}{\tilde{W}(\eta_i, m^*(J_i)|J_i)}\Big), \tag{14}$$

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 (14). More formally

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

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[21].

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.

[21] An alternative would be to base the stopping rule on a formal statistical test of the null hypothesis that $\mathcal{T} = 0$. We did not proceed in this way for two reasons. First the fact that by increasing the number of simulation draws we are free to increase the power of any given alternative to one, suggests that to proceed in this way we would want to formalize the tradeoff between size, power, and the number of simulation draws. Second it seems to us that rather then testing against a null of $\mathcal{T} = 0$ we would want a test with a null which allowed for an acceptable level of error (this even if the only reason for allowing for such errors is the imprecision of the computer's calculations). We note that at the cost of increasing compute time one could produce an estimate of the variance of $\mathcal{T}$ in several different ways (e.g. simply repeat the procedure used to generate the test statistic many times).

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).[22] They note that, at least formally, their stochastic algorithm does away with all aspects of the curse of dimensionality but that in their suggestion for computating a test statistic. Accordingly as they increased the dimension of the state space in their examples 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, thereby eliminating the curse of dimensionality entirely.

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 not 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 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. This implies, however, that 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

[22] 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 (such as Nash conditional on information sets), the computational advantages of the stochastic algorithm would be even greater for then the test would have to be augmented to compute consistent posteriors for each agent active at each point in the state space. These computational difficulties are part of epistemic reasons for thinking that simpler equilibrium concepts, like the one proposed in this paper, might provide a better approximation to actual behavior in applied work.

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 market design would allow generators to make super-normal profits during periods of high demand. In particular the worry is that the twin facts that currently electricity is not storable and has extremely inelastic demand might lead to sharp price increases in periods of high demand (for a review of the literature on price hikes and an empirical analysis of their sources in California during 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 are difficult to incorporate into an equilibrium analysis because, as many authors have noted, they are endogenous.[23]

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, an equilibrium framework for analyzing maintenance decisions requires a dynamic game with strategic interaction. To the best of our knowledge maintenance decisions of electric

[23]There has, however, been an extensive empirical literature on when firms do maintenance (see, for e.g. Harvey, Hogan and Schatzki, 2004, and the literature reviewed their). Of particular interest are empirical investigations of the co-ordination of maintenance decsions, see, for e.g., Wolak and Patrick 2001.

utilities have not been analyzed within such a framework to date. Here we provide a simple example that does endogenizes maintenance decisions, and then ask how asymmetric information effects the results.

Overview of the Model. In our 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 model with asymmetric information the firm knows the cost position of its own generators, but not those of its competitors.

Firms can hold their generators off the market for a single period and do maintenance. Whether they do so 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. In periods when a generator is operated its costs are incremented by a stochastic shock. 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 generators, 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 then compute an AME for each one of them, and compare the results.

Social Planner Problem. To facilitate efficiency comparisons we also present the results generated by the same primitives when maintenance decisions are made by a social planner with full information. The planner maximizes the sum of the discounted value of consumer surplus and net cash flows to the firms. Since the social planner problem is a single agent problem, it was computed using a standard contraction mapping[24].

4.1 Details and Parameterization of The Model.

Firm B has three generators at its disposal. Each of them can produce up to 25 megawatts of electricity at a constant marginal cost which depends on their cost state ($mc_B(\omega)$) and can produce higher levels of electricity at increasing marginal cost. Firm S has four generators at its disposal each of which can produce 15 megawatts of electricity at a constant marginal cost which depends on their cost state ($mc_S(\omega)$) and higher levels at increasing marginal cost. Hence, the marginal cost function of a generator of type $k \in \{B, S\}$ is as follows:

[24]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).

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

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

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

$$\begin{aligned} MC_k(\omega) &= mc_k(\omega) & q < \overline{q}_k \\ &= mc_k(\omega) + \beta(q - \overline{q}_k) & q \geq \overline{q}_k \end{aligned}$$

where $\overline{q}_B = 25$ and $\overline{q}_S = 15$ and the slope parameter $\beta = 10$. For a given level of production, 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 firm 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 identical to its highest marginal cost at which it can operate. The market supply curve is obtained by the horizontal summation of the individual supply curves. For the parameter values indicated in table 1, if firm B bids in N_b number of generators and firm S bids in N_s number of generators, the resultant market supply curve is:

$$Q^{ss}(N_b, N_s) = \begin{cases} 0 & p < 100 \\ 25N_b + (\frac{p-100}{\beta})N_b & 100 \leq p < 200 \\ 25N_b + (\frac{p-100}{\beta})N_b + 15N_s + (\frac{p-200}{\beta})N_b & p \geq 200 \end{cases}$$

The firms horizontally sum the bids of the the generators they decide to bid and submit the resultant bid function to the market maker. The market maker runs a uniform pric auction; it horizontally sums the two firms' bid functions and intersects the resultant aggregate supply curve with the demand curve. This determines the price per megawatt hour and the quantities the two firms are told to produce. Each firm allocates its production across generators so as to minimize its cost; i.e. the generator with the lowest cost state (ω) gets allocated quantity first to a point that the upward sloping part of its marginal cost curve reaches the next generator's minimum marginal cost and so forth.

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 can 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. 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 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

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

where m_i is the bid function which is the highest marginal cost curve of each type of generator. 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.[25] The cost of that maintenance is denoted by cm_i.

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

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

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

where $p(m_{1,t}, m_{2,t}, d_t)$ is the market clearing price, $y_{i,t}(m_{B,t}, m_{S,t}, d_t)$ is the output alocated by the market maker to the i^{th} firm, $I\{\cdot\}$ 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 $\omega_{i,j,t}$, and $\sum_j y_{i,j,t} = y_{i,t}$.

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 we consider. 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 are changes in m_i. We want to point out, however, that these 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.

[25] 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.

4.2 Alternative Informational Assumptions for the Assymmetric 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 (since all generators must do maintenance at least once every six periods, $\tau \leq 5$).

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 (since full revelation occurs every sixth period, 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.[26]

4.3 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 the; (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, for profits, were guaranteed to be higher than their true equilibrium values, and for continuation values, were likely to be higher. 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 given 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

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

The memory was structured first by public information, and then for each given public information node, by the private information of each agent. We

[26]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 that of 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 and computational burdens generated by the different assumptions have to be analyzed numerically.

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 points 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 $\mathcal{L}^2(\mathcal{P}_{\mathcal{R}})$ norm in the percentage deviation between the simulated and estimated values; roughly the weighted average of this deviation where the weights are the number of visits to the point. Had we used the test to determine the stopping iteration and stopped the algorithm whenever the test statistic 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 one minus the test statistic. It shows a rapid fall until about 130 million iterations, and a test statistic that remains essentially unchanged after 150 million (at a value of about .9975).

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 time per one hundred million iterations, each of which includes the test time, is reported in table 2.[27] The differences in compute times across

[27] 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.

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 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, more interestingly, that of the FI model. So if we limit our attention to the recurrent classes the computational demands of the AME AI model are *similar* to those from the FI model. Since numerical analysis of full information dynamic games is often used in applied research, if similar results were true for other environments they would imply that AI models could also be used.

After computing policies we ran a one million iteration simulation from the same initial condition for each of our models. The output from these runs are used for the remainder of the paper. Table 3 compares summary statistics from the three AI models. All summary statistics are virtually identical across the three models (and this was also true for the statistics on policies we consider below). Even the finite history τ model seems to provide enough discrimination between states to approximate the results in the model with full periodic revelation of information. Though these results may be a function of our particular parameterization, they do suggest that, at least for some problems, models with bounded rationality do quite well in approximating unconstrained equilibrium behavior. Since the differences between the AI models are so small, the remainder of the paper only presents results from one of them (the full revelation model).

4.4 Numerical Results.

The output of the algorithm includes a recurrent class of states as well as strategies, costs (both operational and maintenance), profits, and consumer welfare at those states. Here we focus on the maintenance decisions. We begin with the social planner reference point.

The Social Planner Problem. The solution to the social planner problem provides a basis for understanding the logic underlying efficient mainte-

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

nance decisions for our parameterization. Recall that there is significantly less demand on weekends than on weekdays. Table 4 presents average 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 Sundays. 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 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 Duopoly with FI. When there is full information the average 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). 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;

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

about 45% of the time there are two of each type of generator operating, and about 26% of the time three of each type of generator is operating. This leads to over a third of the shutdown decisions for each type of generator occuring when the generator is at one of the two lowest costs states (under 1% of the planner's shutdowns are 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.

The Duopoly with Asymmetric Information. Perhaps the most striking finding in Table 5 is that there is so much less maintenance in the AI than in the FI equilibrium. Indeed our "maintenance frequency" summary statistics from the AI duopoly are much closer to those from the social planner solution than to those from the FI duopoly. The firm with the big generators actually does less maintenance in the AI duopoly then the social planner does, and though the firm with the small generators does do more maintenance, it does about 25% less than the small firm does in the FI duopoly. 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; a fact which contrasts sharply with the behavior of the firms in the FI equilibrium.

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

* Average number of days between maintenance decisions.

Comparing Policies. We wanted to see whether the difference between the AI and FI maintenance decisions was a result of the differences in the states visited in the two equilibria, or to differences in the behavior at a given state. So we took their invariant distributions and asked whether the bids generated by the *static* Nash full information equilibrium in those states differed. They did not differ much.[28] Consequently we focus on reasons for differences in maintenance behavior at a given state.

It seems then that the difference in maintenance behavior is a result of the differences in equilibrium play across states with similar static withdrawl incentives. The incentive to bid an additioanl generator at a given state, say (m_j^o, m_{-j}^o), is given by the expectation of

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

conditional on firm j's information set. If this increment is the (discrete analogue) of convex in m_{-j}^o, then firm j will have a greater incentive to bid an

[28]The static Nash solution for the average number of small generators operating in a day was 2.85 from both invariant measures 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. We note that when computing these results in cases where there were multiple equilbria to the static game, the numbers were computed by averaging over those equilibria.

additional generator in the AI then in the FI equilibrium. We computed the convexity of this difference at each point in the intersection of the recurrent classes from the two equilibira. The average of this statistic was convex, and quite strikingly so. I.e. the incremental profitability of bidding an additional generator increases in the perceived variance in the bids of the firm's competitor. That is, at least in our specification, asymmetric information generates an added incentive to bid generators into the market. This alleviates the incentive to withdraw generators that results from the fact that if a firm withdraws a generator prices are likely to rise and that firm earns more from the electricity produced by the generators it does bid.

The fact that the overall maintenance frequency in the AI equilibrium is similar to the social planner's mainteance frequence, does not, however imply that the *distribution* of maintenance dates in the AI equilibrium is optimal. In particular the AI equilibrium generates more shutdowns on weekends than on weekdays, just the opposite of the social planner. Moreover the equilibrium with asymmetric information sometimes incentivizes the firms to shut down the "wrong" generators; i.e. to shut down generators with lower cost states than those of the other generators that it operates (this is particularly true of the firm with the small generator, and it never occurs in any of the other regimes).

To illuminate this latter point, 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. Note that if a firm's competitor did not change its future bids in response to the firm's own bid, that firm would always do better by shutting down the highest cost generator. Consequently a firm will only not withdraw its most costly generator because a shut down of a higher cost generator induces the firm's competitor to change its bids in the coming period in a way that favors the given firm; i.e. because of the signalling value of the maintenance decision. Indeed when the firm did not shut down the generator with the highest cost it was always the case that 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. I.e. by shutting down the low cost generator the firm insured that the generators that it will operate next period will have operated a longer period of time since their last maintenance decision; a signal of a higher probability of shut down in subsequent periods.

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

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

Consumer Surplus and Profits. Both our duopoly equilibria, that with asymmetric and that with full information, generate a total surplus which are not too different from that generated by the social planner. The full information equilibrium generates a total suplus just 6.5% less than that of the social planner, while the asymmetric information equilibrium generates a surplus which is only 4% less than the social planner. As suggested by the policy comparison, the AI equilibrium generates more consumer surplus but less profits than the FI equilibrium. So the FI equilibrium does do worse when we consider only consumer surplus as it generates only 88-89% of that generated by the planner while the AI equilibrium generates 92-93% of the planner's consumer surplus.

At least two other points are worthy of note. Our excercise only allows for differential demand on weekdays and weekends (not, for example, by time of day). Interestingly even in the solution to the social planner problem we see rather dramatic price effects of this differential demand. Moreover both the

FI and the AI equilibria magnify this difference (though the AI equilibria has lower prices in both demand states). This suggests that institutions which change the pattern of withholding generators are unlikely to have much of an effect on price volatility; to do away with price volatility we will have to find ways to smooth out demand. Also note that it is the firm with the small generators' who gains the most from moving to full information. When there is full information the firm with small generators produces a higher fraction of the output on the lucrative weekdays (58% vs 50%). That is the firm with the low startup cost is better able to adapt to the additional information available in the FI equilibrium.

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 policies which satisfy those policies for 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. Hopefully our framework will enable a more in depth analysis of some of them: those that are well approximated as games that visit a finite set of states repeatedly.

Appendix: Letting η transitions depend on m_{-i}

This appendix provides a generalization which allows the distribution of transitions in one firm's payoff relevant random variables to depend on the actions of its competitors. Formally we change the model to allow for the more general family of η distributions $\mathcal{PG}_\eta$ defined in equation (3) rather than the family $\mathcal{P}_\eta$ in equation (4). This requires us to modify both the definition of an Applied Markov Equilibrium and the computational algorithm.

Definition of Applied Markov Equilibrium. Recall that the definition of equilibrium requires three primitives; (i) a subset $\mathcal{R} \subset \mathcal{S}$; (ii) strategies $(x^*(J_i), m^*(J_i))$ for every J_i which is a component of any $s \in \mathcal{S}$, and (iii) evaluations $W(\eta, m|J_i)$, for each $(\eta, m) \in \Omega(\eta) \times \mathcal{M}$ and every J_i which is a component of any $s \in \mathcal{S}$. The conditions on the policies and evaluations (conditions C2 and C3) change when we allow for the generality in the family $\mathcal{PG}$. Condition C1 ($\mathcal{R}$ is a recurrent class of the process generated by W) does not change. We call the new conditions C2* and C3*.

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 $W(\cdot)$ and the equilibrium play of competitors, that is $(x^*(J_i), m^*(J_i))$ solve

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

where

$$p^e(\eta_i|\ m, x, J_i) \equiv \sum_{J_{-i}} p\ (\eta_i|\ m, x, m^*(J_{-i}), \omega_i) p^e(J_{-i}|J_i),$$

while

$$p^e(J_{-i}|J_i) \equiv \frac{p^e(J_{-i}, J_i)}{p^e(J_i)}.$$

C3*: Consistency of values on $\mathcal{R}$. For every $s \in \mathcal{R}$ let

$$\eta(m^*(s), x^*(s), \omega(s)) \equiv \Big\{ \eta : \Pi_{i=1}^n p\ (\eta_i|m^*(s), x_i^*, \omega_i) > 0 \Big\} \in \Omega(\eta)^n$$

be the set of η vectors that have positive probability when equilibrium strategies are played. Then for every η_i which is a component of an $\eta \in \eta(x^*(s), m^*(s), \omega(s))$ and every $J_i \in s$

$$W(\eta, m^*|J_i) = \pi^E(J_i) \quad + \beta \sum_{J_i'} \Big\{ \sum_{\tilde{\eta}_i} W(\tilde{\eta}_i, m^*(J_i')|J_i') p^e(\tilde{\eta}_i|J_i') \Big\} p^e(J_i'|J_i, \eta_i),$$

where

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

while

$$p^e(\eta_i|J_i) \equiv p^e(\eta_i|m^*(J_i), x^*(J_i), J_i)$$

and

$$p^e(J_i'|J_i, \eta_i) \equiv p^e(J_i', \eta_i|\ J_i)/p^e(\eta_i|J_i). \quad \spadesuit$$

When we allow for the added generality in $\mathcal{PG}_\eta$ the likelihood of the alternative possible $W(\eta, \cdot|\cdot)$ depends on the actions of a firm's competitors. As a result the difference between condition C2* and C2 is that C2* has to insure that the distribution of η the agent uses when chosing its optimal (m, x) is consistent with the distribution of the equilibrium actions of the agent's compeitors.

Condition C3* differs from C3 in two ways. First the set of η_i which are visited with positive probability given equilibium strategies now depends on the actions of the firm's competitors as well as on its own action. This leads to a difference in the definition the set of points at which we must check whether the associated $W(\cdot)$ evaluations are consistent with the primitives of the problem and equilibrium play (of the set $\eta(x^*(s), m^*(s), \omega(s))$.). Second at each of those points the W values must be consistent with the fact that the distribution of η now also depends on the equilibrium actions of the firm's competitors.

Changes in The Computational Algorithm. In the capital accumulation games we focus on in the body of the paper the distribution of η is a primitive. Once we allow those distributions to be partially determined by the actions of competitors as in $\mathcal{PG}_\eta$ the algorithm (and the agent) must learn what those distributions are. So each iteration now is defined by (and has in memory) its location, our $L^k = (J_1^k, \ldots J_{n(k)}^k) \in \mathcal{S}$, and *both*

- the W^k evaluations, and
- the probabilities for competitors play at those states or p^k, where the components of this vector provide perceived probabilities of m_{-i} given J_i.

Often one can simplify and keep a smaller dimensional distribution then that of m_{-i} in memory. For example in the dynamic auction problem discussed in the text where only the highest bid is revealed it would suffice to carry along the distribution of the highest bid conditional on J_i.

The algorithm iteratively updates L^k and (W^k, p^k). The updates for the location are done as before except that; (i) the policies are chosen to solve condition C2* (instead of C2) after substituting W^k for W and $p^k(m_{-i}|J_i^k)$ for $p^e(J_{-i}|J_i^k)$ in that condition, and (ii) the random draws on η conditional on those policies are taken from the family $\mathcal{PG}_\eta$ rather than from $\mathcal{P}_\eta$. To update the W^k we need "perceived realizations" of the value of the outcome. These are now given by

$$V^{k+1}(J_i^k) = \pi(\omega_i^k, \omega_{-i}^k, m_i^{k+1}, m_{-i}^{k+1}, x_i^{k+1}, d^k) \ + \tag{15}$$

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

where

$$p^k(\eta| \ m, x, J_i^{k+1}) \equiv \sum_{m_{-i}} p \ (\eta_i| \ m, x, m_{-i}, \omega_i) p^k(m_{-i}|J_i^{k+1}).$$

The update for W is done as in equation (11) except that we use the definition for $V^{k+1}(J_i^k)$ in equation (15) rather than that in (10).

To update p^k set

$$p^{k+1}(m_{-i}|J_i^k) - p^k(m_{-i}|J_i^k) = \frac{1}{A\Big(h^k(J_i^k)\Big)} \Big[\{m_{-i}^{k+1} = m_{-i}\} - p^k(m_{-i}|J_i^k) \Big], \tag{16}$$

where the curly brackets denotes an indicator function which takes the value of one when the condition inside it is satisfied and zero elsewhere.

That concludes the iteration. The test for equilibrium is the same as that in the text.

References

Athey, S., and K. Bagwell, 2008, "Collusion with Persistent Cost Shocks," *Econometrica*, May 2008, 76 (3), 493-540.

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

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

Beresteanu A. and P. Ellickson (2006) *The Dynamics of Retail Oligopoly*, *mimeo* Department of Economics, Duke Univsersity.

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

Cole, H. and N. Kocherlakota (2001) "Dynamic Games with Hidden Actions and Hidden States", *Journal of Economic Theory.*

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

Doraszelski, U. and A. Pakes (2008), "A Framework for Applied Dynamic Analysis in I.O.," Chapter 33, pp. 2183-2262, in the *The Handbook of Industrial Organization*, M. Armstrong and R. Porter ed.s.

Doraszelski, U. and M. Satterthwaite (2007), "Computable Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity", *mimeo* Harvard University.

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.

Goettler, R., C. Parlour, and U. Rajan (2005) "Equilibrium in a Dynamic Limit Order Market", *The Journal of Finance*, LX No.5, pp. 2149-2192.

Goettler, R., C. Parlour, and U. Rajan (2008) "Informed Traders and Limit Order Markets", *mimeo*, The University of Chicago, the Graduate School of Business.

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.

Kalai, E. and E. Lehrer, (1993) "Rational Learning Leads to Nash Equilibrium", *Econometrica*, 61 (5), pp. 1019-1045.

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

Pakes A., M. Ostrovsky and S. 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.

Wolak F. and R. Patrick (2001); "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales

Electricity Market" NBER Working Paper 8248. ".

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

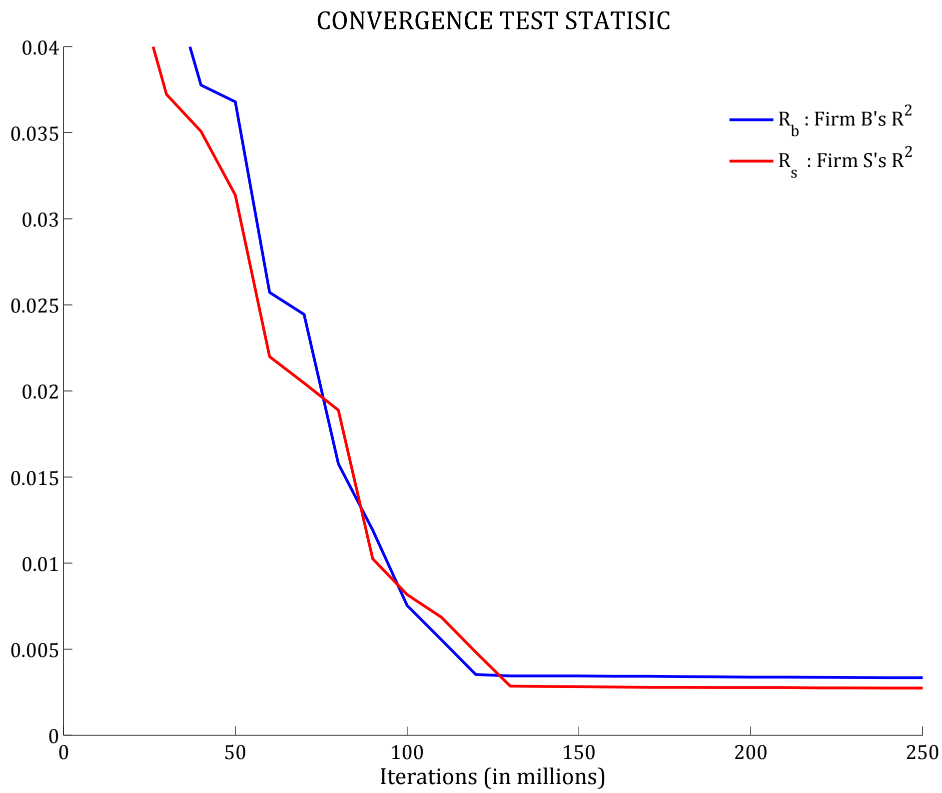
CONVERGENCE TEST STATISIC
0.04
0.035
0.03
0.025
0.02
0.015
0.01
0.005
0
0
50
100
150
200
250
Iterations (in millions)
R_b : Firm B's R^2
R_s : Firm S's R^2

Figure 1: $1 - R^2$