

Dynamic Games with Asymmetric Information: A Framework for Empirical Work*

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

March 18, 2011

Abstract

We develop a framework for the analysis of dynamic oligopolies with persistent sources of asymmetric information that enables applied analysis of situations of empirical importance that have been difficult to deal with. The framework generates policies which are “relatively” easy for agents to use while still being optimal in a meaningful sense, and is amenable to empirical research in that its equilibrium conditions can be tested and equilibrium policies are relatively easy to compute. We conclude with an example that endogenizes the maintenance decisions of electricity generators when the cost 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, Robin Lee, Greg Lewis and Michael Ostrovsky, for helpful comments, and Niyati Ahuja for superb research assistance.

This paper develops a framework for the analysis of dynamic oligopolies with persistent sources of asymmetric information which can be used in a variety of situations which are both of empirical importance and have not been adequately dealt with in prior applied work. These situations include; competition between producers when there is a producer attribute which is unknown to its competitors and serially correlated over time, investment games where the outcome of the investment is unobserved, or repeated auctions for primary products (e.g. timber) where the capacity available to process the quantity acquired by the auction is private information. Less obviously, but probably more empirically important, the framework also allows us to analyze markets in which the decisions of both producers and consumers have dynamic implications, but consumers make decisions with different information sets than producers do. As discussed below, this enables an analysis of dynamics in durable, experience, storeable, and network good industries.

In building the framework we have two goals. First we want a framework which generates policies which are “relatively” easy for agents to use while still being optimal in some meaningful sense of the word. In particular the framework should not *require* the specification and updating of players’ beliefs about their opponents types, as in Perfect Bayesian equilibrium, and should not require agents to retain information that it is impractical for them to acquire. Second we want the framework to be useable by empirical researchers; so its conditions should be defined in terms of observable magnitudes and it should generate policies which can be computed with relative ease (even when there are many underlying variables which impact on the returns to different choices). The twin goals of ease of use to agents and ease of analysis by the applied research work out, perhaps not surprisingly, to have strong complementarities.

To accomplish these tasks we extend the framework in Ericson and Pakes (1995) to allow for asymmetric information.¹ Each agent’s returns in a given period are determined by all agents’ “payoff relevant” state variables and their actions. The payoff relevant random variables of producers would typically include indexes of their cost function, qualities of the goods they market, etc., while in a durable good market those of consumers would include their current holdings of various goods and the household’s own demographic char-

¹Indeed our assumptions nest the generalizations to Ericson and Pakes (1995) reviewed in Doraszelski and Pakes(2008).

acteristics. Neither a player’s “payoff relevant” state variables nor its actions are necessarily observed by other agents. Thus producers might not know either the cost positions or the details of supplier contacts of their competitors, and in the durable goods example neither consumers nor producers need know the entire distribution of holdings crossed with household characteristics (even though this will determine the distribution of future demand and prices).

The fact that not all state variables are observed by all agents and that the unobserved states may be correlated over time implies that variables that are not currently payoff relevant but are related to the unobserved past states of other agents will help predict other agent’s behavior. Consequently they will help predict the returns from a given agent’s current actions. So in addition to payoff relevant state variables agents have “informationally relevant” state variables. For example, in many markets past prices will be known to agents and will contain information on likely future prices.

So the “types” of the agents, which are defined by their state variables, are only partially observed by other agents and evolve over time. In the durable goods example, the joint distribution of household holdings and characteristics will evolve with household purchases, and the distribution of producer costs and goods marketed will evolve with the outcomes of investment decisions. As a result each agent continually changes its perceptions of the likely returns from its own possible actions².

Recall that we wanted our equilibrium concept to be testable. This, in itself, rules out basing these perceptions on Bayesian posteriors, as these posteriors are not observed. Instead we assume that the agents use the outcomes they experienced in past periods that had conditions similar to the conditions the agent is currently faced with to form an estimate of expected returns from the actions they can currently take. Agent’s act so as to maximize the discounted value of future returns implied by these expectations. So in the durable goods example a consumer will know its own demographics and may have kept track of past prices, while the firms might know past sales and prices. Each agent would then chose the action that maximized its estimate of the expected discounted value of its returns conditional on the information at its disposal. We base our equilibrium conditions on the

²Dynamic 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 Kochelakota (2001).

consistency of each agents' estimates with the true expected outcomes.

More formally 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 Experience Based Markov Equilibrium (hereafter, EBE) 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 agent's expected discounted value of net cash flow conditional on the possible actions that agent can take at that state. The conditions we impose on this triple are as follows. The first condition is that the equilibrium policies insure that once we visit a state in our subset we stay within that subset in all future periods, visiting each point in that subset repeatedly; i.e. 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. The final condition is that optimal behavior given these evaluations actually generates expected discounted value of future net cash flows that are consistent with these evaluations on our (recurrent) subset of states.

We show that an equilibrium that is consistent with a given set of primitives can be computed using a simple (reinforcement) learning algorithm. Moreover the equilibrium conditions are testable, and the testing procedure does not require computation of posterior distributions. Neither the iterative procedure which defines the computational algorithm nor the test of the equilibrium conditions have computational burdens which increase at a particular rate as we increase the number of variables which impact on returns; i.e. neither is subject to a curse of dimensionality. At least in principal this should lead to an ability to analyze models which contain many more state variables, and hence are likely to be much more realistic, then could be computed using pointwise Markov Perfect equilibrium concepts³.

³For alternative computational procedures see the review in Doraszelski and Pakes, 2008. Pakes and McGuire, 2001, show that reinforcement learning has significant computational advantages 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 Berenstein and Ellickson, 2006. Goettler, Parlour, and Rajan, 2008, use it to approximate optimal behavior in finance applications. We show that a similar algorithm can be used in games with asymmetric information and provide a test of the equilibrium conditions which is not subject to a curse of dimensionality. The test in the original Pakes and McGuire article was subject to such a curse and it made their algorithm impractical for large problems.

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 the learning rule would not, by itself, restrict behavior without either repeated play or prior information on initial conditions. Also the fact that the equilibrium policies from our model can be learned from past outcomes accentuates the fact that those policies are most 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.

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 subset 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 satisfy inequalities which insure that they are not observed repeatedly. In this sense our notion of equilibrium is akin to the notion of Self Confirming equilibrium, as defined by Fudenberg and Levine (1993) (though our application is to dynamic games). An implication of using the weaker equilibrium conditions is that we might admit more equilibria than the Markov Perfect concept would.

The original Maskin and Tirole (1988) article and the framework for the analysis of dynamic oligopolies in Ericson and Pakes (1995) layed the groundwork for the applied analysis of dynamic oligopolies with symmetric information. This led to a rather large both empirical and numerical literature on an assortment of applied problems (see Benkard, 2004, or Gowrisankaran and Town, 1997, for empirical examples and Doraszelski and Markovich, 2006, or Besanko Doraszelski Kryukov and Satterthwaite, 2010 for examples of numerical analysis). None of these models have allowed for asymmetric information. Our hope is that the introduction of asymmetric information in conjunction with our equilibrium concept helps the analysis in two ways. It enables the applied researcher to use more realistic behavioral assumptions

and hence provide a better approximation to actual behavior, and it simplifies the process of analyzing such equilibria by reducing its computational burden.

As noted this approach comes with its own costs. First it is most likely to provide an adequate approximation to behavior in situations for which there is a relevant history to learn from. Second our equilibrium conditions enhance the possibility for multiple equilibria over more standard notions of equilibria. With additional assumptions one might be able to select out the appropriate equilibria from data on the industry of interest, but there will remain the problem of choosing the equilibria for counterfactual analysis.

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 the solutions to the social planner problem.

The next section describes the primitives of the game. Section 2 provides a definition of, and sufficient conditions for, our notion of an Experience Based Markov Equilibrium. Section 3 provides an algorithm to compute and test for this equilibrium, and section 4 contains our example.

1 Dynamic Oligopolies 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 $n_t \leq \bar{n} < \infty$ (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 costs. The profits of each firm in every period are determined by; their payoff relevant random

⁴Indeed our assumptions nest the generalizations to Ericson and Pakes (1995), and the amendments to it introduced in Doraszelski and Satterthwaite (2010), reviewed in Doraszelski and Pakes(2008). The latter paper also provide more details on the underlying model.

variables, a subset of the actions of all the firms, and a set of variables which are common to all agents and account for common movements in factor costs and demand conditions, 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.

The payoff relevant characteristics of firm i , which will be denoted by $\omega_i \in \Omega_i$, take values on a finite set of points for all i . There will be two types of actions; actions that will be observed by the firm's competitors m_i^o , and those that are unobserved m_i^u . For simplicity we assume that both take values on a finite state space, so $m_i = (m_i^o, m_i^u) \in \mathcal{M}_i$.⁵ Notice that, also for simplicity, we limit ourselves to the case where an agent's actions are either known only to itself (they are "private" information), or to all agents (they are "public" information). For example in an investment game the prices the firm sets are typically observed, but the investments a firm makes in the development of its products may not be. Though both controls could affect current profits and/or the probability distribution of payoff relevant random variables, they need not. A firm might simply decide to disclose information or send a signal of some other form.

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

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

where $\pi(\cdot) : \times_{i=1}^n \Omega_i \times \times_{i=1}^n \mathcal{M}_i \times D \rightarrow R$. $\omega_{i,t}$ evolves over time and its conditional distribution may depend on the actions of all competitors, that is

$$\mathcal{P}_\omega = \{ P_\omega(\cdot | m_i, m_{-i}, \omega); (m_i, m_{-i}) \in \times_{i=1}^n \mathcal{M}_i, \omega \in \Omega \}. \quad (2)$$

Some examples will illustrate the usefulness of this structure.

A familiar special case occurs when the probability distribution of $\omega_{i,t+1}$, or $P_\omega(\cdot | m_i, m_{-i}, \omega)$, does not depend on the actions of a firm's competitors, or m_{-i} . Then we have a "capital accumulation" game. For example in the original Ericson and Pakes (1995) model, m had two components, price and investment, and ω consisted of characteristics of the firm's product

⁵As in Ericson and Pakes (1995), we could have derived the assumption that Ω and $\mathcal{M}$ are bounded sets from more primitive conditions. Also the original version of this paper (which is available on request) included both continuous and discrete controls, where investment was the continuous control. It was not observed by agent's opponents and affected the game only through its impact on the transition probabilities for ω . Appendix 2 provides a way of explicitly incorporating continuous controls into our framework.

and/or its cost function that the firm was investing to improve. Their $\omega_{i,t+1} = \omega_{i,t} + \eta_{i,t} - d_t$, where $\eta_{i,t}$ was a random outcome of the firm's investment whose distribution was determined by $P_\omega(\cdot|m_{i,t}, \omega_{i,t})$, and d_t was determined by aggregate costs or demand conditions.

Now consider a sequence of timber auctions with capacity constraints for processing the harvested timber. Each period there is a new lot up for auction, firms submit bids (a component of our m_i), and the firm that submits the highest bid wins. The quantity of timber on the lot auctioned may be unknown at the time of the auction but is revealed to the firm that wins the lot. The firm's state (our ω_i) is the amount of unharvested timber on the lots the firm owns. Each period each firm decides how much to bid on the current auction (our first component of m_i) and how much of its unharvested capacity to harvest (a second component of m_i which is constrained to be less than ω_i). The timber that is harvested and processed is sold on an international market which has a price which evolves exogenously (our $\{d_t\}$ process), and revenues equal the amount of harvested timber times this price. Then the firm's stock of unharvested in $t + 1$, our $\omega_{i,t+1}$ is $\omega_{i,t}$ minus the harvest during period t plus the amount on lots for which the firm won the auction. The latter, the amount won at auction, depends on $m_{-i,t}$, i.e. the bids of the other firms, as well as on $m_{i,t}$.

Finally consider a market for durable goods. Here we must explicitly consider both consumers and producers. Consumers are differentiated by the type and vintage of the good they own and their characteristics, which jointly define their ω_i , and possibly by information they have access to which might help predict future prices and product qualities. Each period the consumer decides whether or not to buy a new vintage and if so which one (a consumer's m_i); a choice which is a determinant of the evolution of their ω_i . Producers determine the price of the product marketed and the amount to invest in improving their product's quality (the components of the producer's m_i). These decisions are a function of current product quality and its own past sales (both components of the firm's ω_i), as well as other variables which effect the firm's perceptions about demand conditions. Since the price of a firm's competitors' will be a determinant of the firm's sales, this is another example where the evolution of the firm's $\omega_{i,t+1}$ depends on $m_{-i,t}$ as well as on $m_{i,t}$.

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 each agent's strategies are a mapping

from a subset of these variables, in particular from the variables that are observed by the agent *and* are either “payoff” or “informationally” relevant, where these two terms are defined as 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 current profits of at least one of the firms. In terms of equation (1), all components of $(\omega_{i,t}, \omega_{-i,t}, d_t)$ that are observed are 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 $\omega_{j,t}$ which, in turn, will affect $\pi_{i,t}(\cdot)$ through its impact on $m_{j,t}$, player i will generally be able to gain by conditioning its strategy on that variable.⁶

As above 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”). 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 $\omega_{-i,t-1} \in z_{-i,t}$. 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 functions of $J_{i,t}$, i.e.

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

Notice that if $\omega_{-i,t}$ is private information and affects the profits of firm i then we will typically have $\pi_{i,t} \in z_{i,t}$.

We use our examples to illustrate. We can embed asymmetric information into the original Ericson and Pakes (1995) model by assuming that $\omega_{i,t}$ has a product quality and a cost component. Typically quality would be publically observed, but the cost would not be and so becomes part of the firm’s private information. Current and past prices are also part of public information set and contain information on the firms’ likely costs, while investment may be public or private. In the timber auction example, the stock of unharvested

⁶Note that these definitions will imply that an equilibrium in our restricted strategy space will also be an equilibrium in the general history dependent strategy space.

timber is private information, but the winning bids (and possibly all bids), the published characteristics of the lots auctioned, and the marketed quantities of lumber, are public information. In the durable good example the public information is the history of prices, but we need to differentiate between the private information of consumers and that of producers. The private information of consumers consists of the vintage and type of the good it owns and its own characteristics, while the firm's private information includes the quantities it sold in prior periods and typically additional information whose contents will depend on the appropriate institutional structure.

Throughout we only consider games where both $\# \times_{i=1}^n \Omega(\xi)$ and $\# \times_{i=1}^n \Omega(z)$ are *finite*. This will require us to impose restrictions on the structure of informationally relevant random variables, and we come back to a discussion of situations in which these restrictions are appropriate below. To see why we require these restrictions, recall that we want to let agents base decisions on past experience. For the experience to provide an accurate indication of the outcomes of policies we will need a visit a particular state repeatedly; a condition we can only insure when there is a finite state space.

2 Experience Based Equilibrium.

For simplicity we assume all decisions are made simultaneously so there is no subgame that occurs within a period. In particular we assume that at the beginning of each period there is a realization of random variables and players update their information sets. Then the players decide simultaneously on their policies. The extension to sequential decisions within a period (as would be required for the examples with dynamic consumers) are straightforward.

Let s combine the information sets of all agents active in a particular period, that is $s = (J_1, \dots, J_n)$ 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 . We can write s more compactly as $s = (z_1, \dots, z_n, \xi)$. So $\mathcal{S} = \{s : z \in \times_{i=1}^n \Omega_i(z), \xi \in \Omega(\xi), \text{ for } 0 \leq n \leq \bar{n}\}$ lists the possible states of the world.

Firms' strategies in any period are a function of their information sets, so they are a function of a component of that period's s . From equation (2) the strategies of the firms determine the distribution of each firm's information set in the next period, and hence together the firms' strategies determine the

distribution of the next period's s . As a result *any* set of strategies for all agents at each $s \in \mathcal{S}$, together with an initial condition, defines a Markov process on $\mathcal{S}$.

We have assumed that $\mathcal{S}$ is a finite set. As a result each possible sample path of any such Markov process will, in finite time, wander into a subset of the states in $\mathcal{S}$, say $\mathcal{R} \subset \mathcal{S}$, and once in $\mathcal{R}$ will stay within it forever. $\mathcal{R}$ could equal $\mathcal{S}$ but typically will not, as the strategies the agents chose will often insure that some states will not be visited repeatedly, a point we return to below⁷. $\mathcal{R}$ is referred to as a recurrent class of the Markov process as each point in $\mathcal{R}$ will be visited repeatedly.

Note that this implies that the empirical distribution of next period state given any current $s \in \mathcal{R}$ will eventually converge to a distribution which can be constructed from the outcomes at those states. This will also be true of the relevant marginal distributions, for example the joint distribution of the J_i components of s that belong to different firms, or that belong to the same firm in adjacent time periods. We use a superscript e to designate these limiting empirical distributions, so $p^e(J'_i|J_i)$ for $J_i \subset s \in \mathcal{R}$ provides the limit of the empirical frequency that firm i 's next period information set is J'_i conditional on its current information being $J_i \in \mathcal{R}$ and so on⁸.

We now turn to our notion of Experience Based Equilibrium. It is based on the notion that at equilibrium players expected value of the outcomes from their strategies at states which are visited repeatedly are consistent with the actual distribution of outcomes at those states. Accordingly the equilibrium conditions are designed to insure that at such states; (i) strategies are optimal given participants evaluations, and (ii) that these evaluations are consistent with the empirical distribution of outcomes and the primitives of the model.

Notice that this implies that our equilibrium conditions could, at least in principle, be consistently tested⁹. To obtain a consistent test of a condition

⁷Freedman, 1983, provides a precise and elegant explanation of the properties of Markov chains used here. 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 $s \in \mathcal{R}$.

⁸Formally the empirical distribution of transitions in $\mathcal{R}$ will converge to a Markov transition matrix, say $p^{e,T} \equiv \{p^e(s'|s) : (s', s) \in \mathcal{R}^2\}$. Similarly the empirical distribution of visits on $\mathcal{R}$ will converge to an invariant measure, say $p^{e,I} \equiv \{p^e(s) : s \in \mathcal{R}\}$. Both $p^{e,T}$ and $p^{e,I}$ are indexed by a set of policies and a particular choice of a recurrent class associated with those policies. Marginal distributions for components of s are derived from these objects.

⁹We say "in principle" here because this presumes that the researcher doing the testing

at a point we must, at least potentially, observe that point repeatedly. So we could only consistently test for conditions at points in a recurrent class. As we shall see this implies that our conditions are weaker than “traditional” equilibrium conditions. We come back to these issues, and their relationship to past work, after we provide our definition of equilibrium.

Definition: Experience Based Markov Equilibrium. An Experience Based Markov Equilibrium consists of

- A subset $\mathcal{R} \subset \mathcal{S}$;
- Strategies $m^*(J_i)$ 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 the decision m_i , say $W(m_i|J_i)$, for each $m_i \in \mathcal{M}_i$ 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 $\{m^*\}$, 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 $m^*(J_i)$ solves

$$\max_{m_i \in \mathcal{M}_i} W(m_i|J_i)$$

and

C3: Consistency of values on $\mathcal{R}$. Take every J_i which is a component of an $s \in \mathcal{R}$. Then

$$W(m^*(J_i)|J_i) = \pi^E(m^*(J_i), J_i) + \beta \sum_{J'_i} W(m^*(J'_i)|J'_i) p^e(J'_i|J_i),$$

can access the union of the information sets available to the agents playing the game.

where

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

and

$$\left\{ p^e(J'_i|J_i) \equiv \frac{p^e(J'_i, J_i)}{p^e(J_i)} \right\}_{J'_i}, \text{ and } \left\{ p^e(J_{-i}|J_i) \equiv \frac{p^e(J_{-i}, J_i)}{p^e(J_i)} \right\}_{J_{-i}} \spadesuit. \quad (3)$$

Note that the evaluations $\{W(m_i|J_i)\}$ need not be correct for J_i not a component of an $s \in \mathcal{R}$. Nor do we require correctness of the evaluations for the $W(m_i|J_i)$'s associated with points in $\mathcal{R}$ but at policies which differ from those in m_i^* . The only conditions on these evaluations are that choosing an $m_i \neq m_i^*$ would lead to a perceived evaluation which is less than that from the optimal policy (this is insured by condition C2)¹⁰. On the other hand the fact that our equilibrium conditions are limited to conditions on points that are played repeatedly implies that agents are able to learn the values of the outcomes from equilibrium play, and we provide an algorithm that would allow them to form consistent estimates of those outcomes below. Further comments on our equilibrium notion follow.

Beliefs on types. Note also that our conditions are not 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.

Relationship to Self Confirming Equilibria. Experience Based Markov Equilibria, though formulated for dynamic games, is akin to the notion of Self Confirming Equilibria (Fudenberg and Levine, 1993), that has been used in other contexts¹¹. Self Confirming Equilibria weaken the standard Nash

¹⁰The fact that our conditions do not apply to points outside of $\mathcal{R}$ or to $m_i \neq m_i^*$ implies that the conditional probabilities in equation (3) are well defined.

¹¹See also Dekel, Fudenberg and Levine (2004) for an analysis of self confirming equilibrium in games with asymmetric information.

equilibrium conditions. It requires 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 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.

Boundary Points. It is useful to introduce a distinction made 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 boundary points there are actions which lead to outcomes which can not be consistently evaluated by the information generated by equilibrium play.

Multiplicity. Notice that Bayesian Perfect equilibria will satisfy our equilibrium conditions, and typically there will be a multiplicity of such equilibria. Since our experience based equilibrium notion does not restrict perceptions of returns from actions not played repeatedly, it will admit an even greater multiplicity of equilibria. There are at least two ways to select out a subset of these equilibria. One is to impose further conditions on the definition of equilibrium; an alternative which we explore in the next section. As explained their, this requires a game form which enables agents to acquire information on outcomes from non-equilibrium play.

Alternatively (or additionally) if data is available we could use it to restrict the set of equilibria. I.e. if we observe or can estimate a subset of either $\{W(\cdot)\}$ or $\{m^*(\cdot)\}$ we can restrict any subsequent analysis to be consistent with their values. In particular since there are (generically) unique equilibrium strategies associated with any given equilibrium $\{W(\cdot)\}$, if we were able to determine the $\{W(\cdot)\}$ associated with a point (say through ob-

servations on sample paths of profits) we could determine m_i^* that point, and conversely if we know m_i^* at a point we can restrict equilibrium $\{W(\cdot)\}$ at that point. Similarly we can direct the computational algorithm we are about to introduce to compute an equilibria that is consistent with whatever data is observed. On the other hand were we to change a primitive of the model we could not single out the equilibria that is likely to result without further assumptions (though one could analyze likely counterfactual outcomes if one is willing to assume a learning rule and an initial condition; see Lee and Pakes, 2009).

2.1 Restricted Experience Based Equilibria.

Our condition C3 only requires correct evaluations of outcomes from equilibrium actions that are observed repeatedly; i.e. for $W(m_i|J_i)$ at $m_i = m_i^*$ and $J_i \subset s \in \mathcal{R}$. There are circumstances when imposing restrictions on equilibrium evaluations of actions off the equilibrium path for states that are observed repeatedly, that is at $m_i \neq m_i^*$ for $J_i \subset s \in \mathcal{R}$, might be natural, and this subsection explores them.

Barring compensating errors, for agents to have correct evaluations of outcomes from an $m_i \neq m_i^*$ they will need to know; (i) expected profits and the distribution of future states that result from playing m_i , and (ii) the continuation values from the states that have positive probability when m_i is chosen. Whether or not agents can obtain the information required to compute expected profits and the distribution of future states when an $m_i \neq m_i^*$ is played depends on the details of the game, and we discuss this further below. For now we assume that they can, and consider what this implies for restricting the evaluations of outcomes from non-optimal actions.

Consider the restriction which assumes the analogue of C3 for all $m_i \in \mathcal{M}_i$ at any $J_i \subset s \in \mathcal{R}$. Then, at equilibrium, all outcomes that are in the recurrent class are evaluated in a way that would be consistent with the expected discounted value of returns that the action would yield were all agents (including itself) to continue playing their equilibrium strategies; and this regardless of whether the action that generated the outcome was an equilibrium action. As in an unrestricted EBE outcomes that are not in the recurrent class are evaluated by perceptions which are not required to be consistent with any observed outcome¹². As a result the restricted EBE

¹²We note that there are cases where it would be natural to require outcomes not in the

insures that in equilibrium when agents are at interior points they evaluate all feasible actions in a way that is consistent with expected returns given equilibrium play. However at boundary points only those actions whose outcomes are in the recurrent class with probability one are evaluated in this manner.

Definition: Restricted Experience Based Equilibrium. Let $\pi^E(m_i, J_i)$ be expected profits and $\{p(J'_i|J_i, m_i)\}_{J'_i}$ be the distribution of J' , both conditional on (m_i, J_i) and m_{-i}^* . A strongly restricted EBE requires, in addition to C1 and C2, that

$$W(m_i | J_i) = \pi^E(m_i, J_i) + \beta \sum_{J'_i} W(m^*(J'_i)|J'_i)p(J'_i|J_i, m_i) \quad (4)$$

for all $m_i \in \mathcal{M}_i$ and $J_i \subset s \in \mathcal{R}$.

Notice that there is a sense in which this equilibrium concept imposes a strong restriction on how agent's react to non-equilibrium play by their competitors. To see this recall that the outcome is $J'_i = (\xi', z'_i)$, where ξ' contains new public, and z'_i new private, information. Competitors observe ξ' and ξ . Were an agent to play an $m_i \neq m_i^*$ it may generate a ξ' which is not in the support of the distribution of ξ' generated by (ξ, m_i^*) ¹³. Then if we impose the restrictions in equation (4) we impose constraints on the agent's evaluations of outcomes of actions which the agent's competitors would see as inconsistent with equilibrium play. For the agent to believe such estimates are correct, the agent would have to believe that the competitor's play would not change were the competitor to observe an action off the equilibrium path. We could have assumed that, in equilibrium, agents only need to have correct evaluations for the outcomes of actions that competitor's could see as consistent with equilibrium play; i.e. actions which generate a support for ξ' which is contained in the support ξ' conditional on (ξ, m_i^*) . Then we

recurrent class to be consistent with publically available information on primitives. For example even if a firm never exited from a particular state the agent might know its selloff value, so it might be reasonable to require that the action of exiting to be evaluated at that value. It is straightforward to impose such constraints on the computational algorithm introduced in the next section.

¹³As an example consider the case where m_i is observable. Then were the agent to play $\tilde{m}_i \neq m_i^*$, $\tilde{m}_i$ would be in ξ' and, provided there does not exist a $\tilde{J}_i = (\xi, \tilde{z}_i)$ such that $m^*(\xi, \tilde{z}_i) = \tilde{m}_i$, the support of ξ' given $(\xi, \tilde{m}_i)$ will differ from that given (ξ, m_i^*) .

would only restrict equilibrium beliefs about outcomes from actions that no agent perceives as inconsistent with equilibrium play. We do not pursue this further here, but it would be straightforward to modify the computational algorithm introduced in section 3 to accomodate this definition of a restricted EBE rather than the one in equation (4).

As noted for agents to be able to evaluate actions in a manner consistent with the restricted EBE they must know $\pi^E(m_i, J_i)$ and $\{p(J'_i|J_i, m_i)\}_{J'_i}$ for $m_i \neq m_i^*$ at all $J_i \subset s \in \mathcal{R}$. Sometimes these objects can be computed from the information generated by equilibrium play and/or knowledge of the primitives of the problem, and sometimes they can not. We illustrate with a familar game below. Note, however, that even if agents can access the required information, to evaluate actions in the way assumed in a restricted EBE they will have to incur the cost of storing additional information and making additional computations; a cost we return to in the context of the computational algorithm discussed in the next section.

Assume that $\pi^E(m_i, J_i)$ can be consistently estimated¹⁴, and consider conditions which enable the agent to consistently estimate $p(J'_i|J_i, m_i) \equiv p(\xi'|z'_i, J_i, m_i)p(z'_i|J_i, m_i)$, $\forall J'_i$ in the support of (m_i, J_i) , $m_i \in \mathcal{M}_i$ and $J_i \subset s \in \mathcal{R}$. We begin with a capital accumulation game in which m_i is not observed and only impacts the distribution of ξ through its impact on the distribution of z_i . For specificity sssume z_i represent costs, m_i is unobserved cost reducing investment, and $p(\xi'|z'_i, J_i, m_i) = p(\xi'|z'_i, J_i)$. Assume also that $\{z\}$ evolves as a controled Markov process, so that $p(z'_i|J_i, m_i) = p(z'_i|z_i, m_i)$, and is known from the primitives of the cost reducing process. Though costs are not observed they are a determinant of prices which are, so in this game past prices contain information on current costs.

Then an agent will be able to consistently estimate $p(J'_i|J_i, m_i)$ whenever the agent can consistently estimate $p(\xi'|z'_i, J_i)$. If $p(\xi'|z'_i, J_i) = p(\xi'|J_i)$ for

¹⁴Whether or not $\pi^E(m_i \neq m_i^*, J_i)$ can be consistently estimated depends on the specifics of the problem, but it frequently can be. For a simple example consider an investment game where the profit function is additively separable in the cost of investment or m_i , so that $\pi^E(m_i, J_i) = \pi^E(m_i^*, J_i) + m_i^* - m_i$. If profits are not additively separable in m_i but m_i is observed then it suffices that agents be able to compute profits as a function of (J_i, m_i, m_{-i}) , as in the computational example below and in differentiated product markets in which the source of assymetric information is costs, equilibrium is Nash in prices, and agents know the demand function. In auctions the agent can compute $\pi^E(m_i, J_i)$ if the agent can learn the distribution of the winning bid.

all ξ' , as would be the case if the current period's investment outcome did not affect the current period's prices, then the outcomes from equilibrium play can be used to obtain consistent estimates of the required probabilities. Now say $p(\xi'|z'_i, J_i) \neq p(\xi'|J_i)$ for some ξ' , as would occur if the innovation in costs affected today's prices. Equilibrium play generates $\{p(\xi', z'_i|J_i, m_i^*)\}$, so since $p(z'_i|J_i, m_i)$ is a primitive, the agent can consistently estimate

$$p(\xi'|z'_i, J_i) = \left(\frac{p(\xi', z'_i|J_i, m_i^*)}{p(z'_i|z_i, m_i^*)} \right)$$

for every z' in the support generated by m_i^* . As a result the agent will be able to consistently estimate the distribution of J'_i given $(J_i, m_i \neq m_i^*)$ whenever the support of z'_i given $(z_i, m_i \neq m_i^*)$ is contained in that of (z_i, m_i^*) .

Now consider the same case but allow for an m_i which generates a support for z'_i which is not contained in that generated by m_i^* . An argument analogous to that above will insure that the agent will be able to consistently estimate $p(\xi'|z'_i, J_i)$ only if there is $\tilde{J}_i = (\tilde{\xi}, \tilde{z}_i) \subset s \in \mathcal{R}$ and associated $m^*(\tilde{\xi}, \tilde{z}_i)$, which generate probabilities, $p(\xi'|z'_i, \tilde{J}_i)$, equal to $p(\xi'|z', J_i)$ for every (ξ', z'_i) in the support of (m_i, J_i) . This would occur in our example if current prices were just a function of current costs and the public information, and we could find a $\tilde{J}_i = (\tilde{\xi}, \tilde{z}_i)$, i.e where the public information in $\tilde{J}_i$ is the same as that in J_i , and the support of z'_i given $(m^*(\tilde{J}_i), \tilde{z}_i)$ contains that of (m_i, z_i) .

The conditions when m_i is observed are similar. When m_i is observed then m_i will generally be a component of ξ' , so $p(\xi'|z'_i, J_i, m_i)$ will no longer be independent of m_i (the exception is the case where all private information is revealed, so m_i is redundant). If $m_i \in \xi'$ and $\xi \in \xi'$ (so there is no forgetting), then the agent will be able to consistently estimate the distribution of J'_i given $(J_i, m_i \neq m_i^*)$ only if there is another point in the recurrent class, say $\tilde{J}_i = (\tilde{\xi}, \tilde{z}_i)$ such that $\tilde{\xi} = \xi$, $m^*(\tilde{\xi}, \tilde{z}_i) = m_i$, and the support of z'_i given (z_i, m_i) is contained in that given $(\tilde{z}_i, m^*(\tilde{J}_i))$ ¹⁵.

We show how to compute and test for a restricted EBE in section 3.

¹⁵Similar considerations apply to non capital accumulation games (games where a competitor's play affects the evolution of one's own state), with the added complication that since $p(z'_i|J_i, m_i)$ depends on the competitor's control, we also have to check whether we can form those probabilities for $m_i \neq m_i^*$ from equilibrium play.

2.2 The Finite State Space Condition.

Our framework is restricted to finite state games. We now consider this restriction in more detail. We have already assumed that there was: (i) an upper bound to the number of firms simultaneously active, and (ii) each firm's physical states (our ω) could only take on a finite set of values. These restrictions insure that the payoff relevant random variables are finite dimensional, but they do not guarantee this for the informationally relevant random variables, so optimal strategies could still depend on an infinite history¹⁶. We can insure that the informationally relevant random variables are finite dimensional either; (i) through restrictions on the form of the game, or (ii) by imposing constraints on the cognitive abilities of the decision makers.

One example of a game form which can result in a finite dimensional space for the informationally relevant state variables is when there is periodic simultaneous revelation of all variables which are payoff relevant to all agents. Claim 1 of Appendix 1 show that in this case an equilibrium with strategies restricted to depend only a finite history is an equilibrium to the game with unrestricted strategies. Claim 2 of Appendix 1 shows that there is indeed a restricted EBE for the game with periodic revelation of information. The numerical analysis in section 4 includes an example in which regulation generates such a structure. Periodic revelation of all information can also result from situations in which private information can seep out of firms (say through labor mobility) and will periodically do so for all firms at the same time, and/or when the equilibrium has one state which is visited repeatedly at which equilibrium play reveals the states of all players.

There are other game forms which insure finiteness. One example is when the institutional structure insures that 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

¹⁶The conditions would however insure finiteness in a game with asymmetric information where the sources of asymmetric information are distributed independently over time (as in Bajari, Benkard and Levin, 2007, or Pakes Ostrovsky and Berry, 2007).

any bidder's information set, and hence can not effect bids.¹⁷ Alternatively a combination of assumptions on the functional forms for the primitives of the problem and the form of the interactions in the market that yield finite dimensional sufficient statistics for all unknown variables could also generate our finite state space condition.

A different way to insure finiteness is through bounded cognitive abilities, say through a direct bound on memory (e.g., agents can not remember what occurred more than a finite number of periods prior), or through bounds on complexity, or perceptions. There are a number of reasons why such a restriction may be appealing to empirical researchers. First it might be thought to be a realistic approximation to the actual institutions in the market. Second in most applications the available data is truncated so the researcher does not have too long a history to condition on. Moreover in any given application one could investigate the extent to which policies and/or outcomes depended on particular variables either empirically or computationally.

To illustrate our computational example computes equilibria to finite state games generated by both 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.

3 An Algorithm for Computing an EBE.

This section shows that we can use a reinforcement learning algorithm to compute an Experience Based Equilibrium. As a result our equilibria can be motivated as the outcome of a learning process. In the reinforcement learning algorithm players form expectations on the value that is likely to result from the different actions available to them and choose their actions optimally given those expectations. From a given state those actions, together with realizations of random variables whose distributions are determined by them, lead to a current profit and a new state. Players use this profit together with their expectations of the value they assign to the new state to update their expectation of the continuation values from the starting state. They then proceed to chose an optimal policy for the new state, a policy

¹⁷Formally 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.

which maximizes its expectations of the values from that state. This process continues iteratively.

Note that the players' evaluations at any iteration need not be correct. However we would expect that if policies converge and we visit a point repeatedly we will eventually learn the correct continuation value of the outcomes from the policies at that point. Our computational mimic of this process includes a test of whether our equilibrium conditions, conditions which insure that continuation evaluations are in fact consistent with subsequent play, are satisfied. 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 subsection¹⁸.

We begin with the iterative algorithm for an EBE, then note the modifications required for a restricted EBE, and then move on to the test statistic for both equilibrium concepts. A discussion of the properties of the algorithm, together with its relationship to the previous literature and additional details that can make implementation easier, is deferred until Appendix 2.

The algorithm consists of an iterative procedure and subroutines for calculating initial values and profits. We begin with the iterative procedure. Each iteration, indexed by k , starts with a location which is a state of the game (the information sets of the players) say $L^k = [J_1^k, \dots, J_{n(k)}^k]$, and the objects in memory, say $M^k = \{M^k(J) : J \in \mathcal{J}\}$. The iteration updates both these objects. We start with the updates for an unrestricted EBE, and then come back to how the iterative procedure is modified when computing a restricted EBE. The rule for when to stop the iterations consists of a test of whether the equilibrium conditions defined in the last section are satisfied, and we describe the test immediately after presenting the iterative scheme.

Memory. The elements of $M^k(J)$ specify the objects in memory at iteration k for information set J , and hence the memory requirements of the algorithm. Often there will be more than one way to structure the memory with different ways having different advantages. Here we focus on a simple

¹⁸On the other hand, there are several issues that arise were one to take the learning approach as an approximation to behavior, 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.

structure which will always be available (though not necessarily always be efficient); alternatives are considered in Appendix 2.

$M^k(J_i)$ contains

- a counter, $h^k(J_i)$, which keeps track of the number of times we have visited J_i prior to iteration k , and if $h^k(J_i) > 0$ it contains
- $W^k(m_i|J_i)$ for $m_i \in \mathcal{M}_i, i = 1, \dots, n$.

If $h^k(J_i) = 0$ there is nothing in memory at location J_i . If we require $W(\cdot|J_i)$ at a J_i at which $h^k(J_i) = 0$ we have an initiation procedure which sets $W^k(m_i|J_i) = W^0(m_i|J_i)$. Appendix 2 considers choices of $\{W^0(\cdot)\}$. For now we simply note that high initial values tend to insure that all policies will be explored.

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 $m^k(J_i^k)$ to

$$\max_{m \in \mathcal{M}_i} W^k(m|J_i^k).$$

With this $\{m^k(J_i^k)\}$ 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, $\{m^k(J_i^k)\}$, are then substituted into the conditioning sets for the distributions of the next period's state variables (the distributions in equation 2 for payoff relevant random variables and the update of informationally relevant state variables if the action causes such an update), and they, in conjunction with the information in memory at L^k , determine a distribution for future states (for $\{J_i^{k+1}\}$). A pseudo random number generator is then used to obtain a draw on the next period's payoff relevant states.

Updating. Use $(J_i^k, m^k(J_i^k), \omega_i^{k+1}, d^{k+1})$ 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

$$V^{k+1}(J_i^k) = \pi(\omega_i^k, \omega_{-i}^k, m_i^k, m_{-i}^k, d^k) + \max_{m \in \mathcal{M}_i} W^k(m|J_i^{k+1}) \quad (5)$$

Note that to calculate $V^{k+1}(J_i^k)$ we need to first find and call up the information in memory at locations $\{J_i^{k+1}\}_{i=1}^{n_{k+1}}$.¹⁹ 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 (5) 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, for example

$$W^{k+1}(m|J_i^k) = \frac{1}{h^k(J_i^k)} V^{k+1}(J_i^k) + \left(\frac{h^k(J_i^k) - 1}{h^k(J_i^k)} \right) W^k(m|J_i^k), \quad (6)$$

which makes $W^k(J_i^k)$ the simple average of the $V^r(J_i^r)$ over the iterations at which $J_i^r = J_i^k$. Though use of this simple average will satisfy Robbins and Monroe's (1951) convergence conditions, we will typically be able to improve the precision of our estimates of the $W(\cdot)$ by using a weighting scheme which downweights the early values of $V^r(\cdot)$ as they are estimated with more error than the later values.²⁰

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 $i = 1, \dots, n_k$) 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

¹⁹The burden of the search for these states depends on how the memory is structured, and the efficiency of the alternative possibilities depend on the properties of the example. As a result we come back to this question when discussing the numerical example below.

²⁰One simple, and surprisingly effective, way of doing so is to restart the algorithm using as starting values the values outputted from the first several million draws. The Robbins and Monroe, 1951, article is often considered to have initiated the stochastic approximation literature of which reinforcement learning is a special case. Their conditions on the weighting function are that the sum of the weights of each point visited infinitely often must increase without bound while the sum of the weights squared must remain bounded.

and values the algorithm outputs are equilibrium policies and values. We come back to that test presently.

Updating when computing a restricted EBE. The algorithm just described only updates $W^k(m_i|J_i)$ for $m_i = m^k(J_i^k)$, the policy that is optimal given iteration k 's evaluations. Recall that we only expect correct evaluations from outcomes that are updated repeatedly. Since this algorithm only updates the continuation values of actions actually taken, if the algorithm does converge to an EBE only the continuation values of equilibrium actions are likely to satisfy C3. That is the algorithm is unlikely to provide correct evaluations of outcomes from actions off its equilibrium path, and a restricted EBE requires correct evaluations of some of those outcomes (the outcomes in $\mathcal{R}$).

When we want to compute a restricted EBE we modify the algorithm just described to update all the $\{W^k(m|J_i^k)\}_{m \in \mathcal{M}_i}$, i.e. the continuation values for *all possible* actions from a state whenever that state is reached. This insures that whenever a non-equilibrium action has a possible outcome which is in the recurrent class it will be evaluated correctly provided all recurrent class points are evaluated correctly. To update $W^k(m_i|J_i^k)$ when $m_i \neq m^k(J_i^k)$ we take a random draw from the distribution of outcomes conditional on that m_i , use it and the random draws from the competing agent's *optimal* policies to form what the perceived value realization would have been had the agent implemented policy $m_i \neq m_i^*$ (the analogue of $V^{k+1}(J_i^k|m_i)$ in equation 5), and use it to form $W^{k+1}(m_i|J_i^k)$ (as in equation 6). The rest of the algorithm is as above; in particular we update the location using the draws from the optimal policy. Note that this increases both the computational and memory requirements of the algorithm significantly. Both the computational burden and the memory requirements at a point, L^k , grows by a factor of $\times_{i=1}^{n_k} \#\mathcal{M}_i/n_k$.

3.1 Testing for an EBE.

Assume we have a W vector in memory at some iteration of the algorithm, say $W^k = \tilde{W}$, and we want to test whether $\tilde{W}$ generates an EBE on a recurrent subset of $\mathcal{S}$. To perform the test we need to check our equilibrium conditions and this requires: (i) a candidate for a recurrent subset determined by $\tilde{W}$, say $\mathcal{R}(\tilde{W})$, and checks for both (ii) the optimality of policies and (iii) the consistency of $\tilde{W}$, on $\mathcal{R}(\tilde{W})$.

To obtain a candidate for $\mathcal{R}(\tilde{W})$, start at any s^0 and use the policies implied by $\tilde{W}$ 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 $j = J_2$. Provided J_1 and J_2 grow large $\mathcal{R}$ will become a recurrent class of the process generated by $\tilde{W}$. In practice to determine whether any finite (J_1, J_2) are large enough, one generates a second sample path starting at J_2 and continuing for another $J_2 - J_1$ iterations. We then check to see that the set of points visited on the second sample path are the same as those in $\mathcal{R}(J_1, J_2, \cdot)$.

The second equilibrium condition specifies that the policies must be optimal given $\tilde{W}$. This is satisfied by construction as we chose the policies that maximize $\tilde{W}(m_i|J_i)$ at each J_i .

To check the third equilibrium condition we have to check for the consistency of $\tilde{W}$ with outcomes from the policies generated by $\tilde{W}$ on the points in $\mathcal{R}$. Formally we have to check for the equality in

$$\tilde{W}(m_i^*|J_i) = \pi^E(m_i^*, J_i) + \beta \sum_{J'_i} \left\{ \tilde{W}(m^*(J'_i)|J'_i) \right\} p^e(J'_i|J_i).$$

In principle we could check this by direct summation for the points in $\mathcal{R}$. However this is computationally burdensome, and the burden increases exponentially with the number of possible states (generating a curse of dimensionality). So proceeding in this way would limit the types of empirical problems that could be analyzed.

A far less burdensome alternative, and one that does not involve a curse of dimensionality, is to use simulated sample paths for the test. To do this we start at an $s_0 \in \mathcal{R}$ and forward simulate. Each time we visit a state we compute perceived values, the $V^{k+1}(\cdot)$ in equation (5), for each J_i at that state, and keep track of the average and the sample variance of those simulated perceived values across visits to the same state, say

$$\left\{ \left(\hat{\mu}(W(m^*(J_i)|J_i)), \hat{\sigma}^2(W(m^*(J_i)|J_i)) \right) \right\}_{J_i \subset \mathcal{R}, s \in \mathcal{R}}.$$

An estimate of the mean square error of $\hat{\mu}(\cdot)$ as an estimate of $\tilde{W}(\cdot)$ can be computed as i.e. $(\hat{\mu}(\cdot) - \tilde{W})^2$. The difference between this mean square error and the sampling variance, or $\hat{\sigma}^2(W(m^*(J_i)|J_i))$, is an unbiased estimate of the bias squared of $\hat{\mu}(\cdot)$ as an estimate of $\tilde{W}(\cdot)$. We base our test of the third EBE condition on these bias estimates.

More formally if we let $E(\cdot)$ take expectations over simulated random draws, l index information sets, and do all computations as percentages for

each $\tilde{W}_l(\cdot)$ value, the expectation of our estimate of the percentage mean square of $\hat{\mu}(W_l)$ as an estimate of $\tilde{W}_l$ is

$$MSE_l \equiv E[\hat{MSE}_l] \equiv E\left(\frac{\hat{\mu}(W_l) - \tilde{W}_l}{\tilde{W}_l}\right)^2 = \quad (7)$$

$$E\left(\frac{\hat{\mu}(W_l) - E[\hat{\mu}(W_l)]}{\tilde{W}_l}\right)^2 + \left(\frac{E[\hat{\mu}(W_l)] - \tilde{W}_l}{\tilde{W}_l}\right)^2 \equiv \sigma_l^2 + (Bias_l)^2$$

Let $(\hat{MSE}_s, \sigma_s^2, (Bias_s)^2)$ be the average of $(\hat{MSE}_l, \sigma_l^2, (Bias_l)^2)$ over the information sets (the l) of the agents active at state s , and $\hat{\sigma}_s^2$ be the analogous average of $\hat{\sigma}^2(W_l)/\tilde{W}_l^2$. Then since $\hat{\sigma}_s^2$ is an unbiased estimate of σ_s^2 , the law of large numbers insures that an average of the $\hat{\sigma}_s^2$ at different s converges to the same average of σ_s^2 . Let h_s be the number of times we visit point s . We use as our test statistic, say $\mathcal{T}$, an h_s weighted average the difference between the estimates of the mean square and that of the variance, and if $\rightarrow$ indicates (almost sure) convergence, the argument above implies that

$$\mathcal{T} \equiv \sum_s h_s \hat{MSE}_s - \sum_s h_s \hat{\sigma}_s^2 \rightarrow \sum_s h_s (Bias_s)^2, \quad (8)$$

a weighted average of the sum of squares of the percentage bias. If $\mathcal{T}$ is sufficiently small we stop the algorithm; otherwise we continue ²¹.

Testing for a restricted EBE. Our test for a restricted EBE is similar except that in the restricted case we simulate the mean and the variance of outcomes for every $m_i \in \mathcal{M}_i$ for each information set l , say $(\hat{\mu}_{m_i,l}, \hat{\sigma}_{m_i,l}^2)$, for each $J_l \subset s$ and $s \in \mathcal{R}$. We then use the analogue of equation (7) to derive estimates of $\{\hat{MSE}_{l,m_i}\}$ and average over $m_i \in \mathcal{M}_i$ to obtain new estimates of $(\hat{MSE}_l, \hat{\sigma}_l^2)$. The test statistic is obtained by substituting these new estimates into the formula for $\mathcal{T}$ in equation (8) above, and will be labeled $\mathcal{T}_R$.

²¹Formally $\mathcal{T}$ is an $L^2(\mathcal{P}_{\mathcal{R}})$ norm in the percentage bias, where $\mathcal{P}_{\mathcal{R}}$ is the invariant measure associated with $(\mathcal{R}, \tilde{W})$. Appendix 2 comments on alternative possible testing procedures, some of which may be more powerful than the test provided here.

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 possibility of “forced”, in contrast 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.²²

Since the benefits from incurring 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 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.

²²There 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 there). Of particular interest are empirical investigations of the co-ordination of maintenance decisions, see, for e.g., Patrick and Wolak, 1997.

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, and hence on the timing of Firm B's prior maintenance decisions. 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 EBE 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 that knows the cost states of all generators. 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 mapping²³.

4.1 Details and Parameterization of The Model.

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.

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

²³The 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 asymmetric information duopoly (see section 3 and the details below).

marginal cost. Hence, the marginal cost function of a generator of type $k \in \{B, S\}$ is as follows:

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

where $\bar{q}_B = 25$ and $\bar{q}_S = 15$ 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:

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

The market maker runs a uniform price auction; it horizontally sums the generators' 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. The market maker then allocates production across generators in accordance with the bid functions and the equilibrium price.

The demand curve is log-linear

$$\log(Q)^{MD} = D_d - \alpha \log(P),$$

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 generator is not operated in this period it does maintenance and at the beginning of the next period can be operated at the low cost base state ($\omega = 0$). If the generators is operated the state of the generator stochastically

decays. Formally if $\omega_{i,j,t} \in \Omega = \{0, 1, \dots, 4\}$ 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

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

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 $\omega_{i,t} = \{\omega_{i,j,t}; j = 1 \dots n_i\} \in \Omega^{n_i}$, 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 $i \in \{S, B\}$ is a choice of

$$m_i = [m_{1,i}, \dots, m_{n_i,i}] : J_i \rightarrow (0, m_i)^{n_i} \equiv M_i,$$

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.²⁴ The cost of that maintenance is denoted by cm_i .

If $p(m_{1,t}, m_{2,t}, d_t)$ is the market clearing price while $y_{i,j,t}(m_{B,t}, m_{S,t}, d_t)$ is the output allocated by the market maker to the j^{th} generator of the i^{th} firm, the firm's profits ($\pi_i(\cdot)$) are

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

²⁴In none of our runs was this constraint binding more than in .29% of the cases, and in most cases it never bound at all.

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

where; $I\{\cdot\}$ is the indicator function which is one if the condition inside the brackets is satisfied and zero elsewhere, $c(\omega_{i,j,t}, y_{i,j,t}(\cdot))$ is the cost of producing output $y_{i,j,t}$ at a generator whose cost state is given by $\omega_{i,j,t}$, and $cm_{j,i}$ is the cost of maintenance (our “investment”).

Note. We now go on to describe the different public information sets that we allow the firm to condition on in the three asymmetric information models we consider. All three, when combined with our functional form and behavioral assumptions, produce quite simple special cases of our general model. In particular they imply that: (i) the only 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, both that it is straightforward to add realism to the model, and that this simple example is unlikely to generate an adequate approximation to any real electricity market. The current assumption were chosen to keep the model transparent and to make it easier to isolate the impact of asymmetric information on equilibrium behavior.

4.2 Alternative Informational Assumptions for the Asymmetric Information Model.

As noted the public information that is informationally relevant 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 assumptions each of which have the effect of insuring finiteness, and then compare the results they generate and their computational properties.

All three asymmetric information (henceforth, AsI) 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. One is the case of periodic full revelation of information; it is assumed that a regulator inspects all generators during every fifth period and announces the states of all generators just before the sixth period. In this case we know

that if one agent uses strategies that depend only on the information it has accumulated since the states of all generators were revealed, the other agent can do no better than doing so also. The other two cases restrict the memory used in the first case; in one a firm partitions the history it uses more finely than in the other. In these cases it may well be that the agents would have profitable deviations if we allowed them to condition their strategies on more information.

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

1. 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).
2. In finite history " m " the public information is just the maintenance decisions made in each of the last five periods on each generator.
3. In finite history " τ " the public information is only the time since the last maintenance decision of each generator (since all generators must do maintenance at least once every six periods, $\tau \leq 5$).

The information kept in memory in each period in the third 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 first 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 opportunities when employing the equilibrium strategies of the last 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.²⁵

²⁵However there is no necessary relationship between the size of the recurrent classes in the alternative models, and as a result no necessary relationship between either the computational burdens or the memory requirements of those models. The memory requirements and computational burdens generated by the different assumptions have to be analyzed numerically.

4.3 Computational Details.

The EBE 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. These include; (i) starting values for the $W(\cdot|\cdot)$'s and the $\pi^E(\cdot|\cdot)$, (ii) information storage procedures, and (iii) the testing procedure.

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, that we were quite sure would be higher. Our initial 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 initial 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 and no depreciation, that is

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

The memory was structured first by public information, and then for each given public information node, by the private information of each agent. We 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 our test statistic is a weighted average of the squared percentage bias in our estimates of the continuation values, 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 periodic 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).

4.4 Computational Properties of the Results.

The computational properties we focus on are; (i) the compute times (ii) the sizes of the recurrent class and hence the memory requirements of both computers and, perhaps more importantly, of agents formulating strategies, and (iii) comparisons of the outputs from the asymmetric information models which bounded memory to those that did not.

Table 2: **Computational Comparisons.**

	AsI; Finite Hist. τ	AsI; Finite Hist. m	AsI; 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 (including the test time) and the sizes of the recurrent classes are reported in table 2.²⁶ The differences in compute times across models roughly reflect the differences in the size of the recurrent class from the different specifications, as this determines the search time required to bring up and store information, and none of them seem prohibitive.

The sizes of the recurrent classes are notable for several reasons. First the recurrent class in the model with periodic full revelation of information

²⁶All 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.

is about the same size as that of the full information model. This is despite the fact that the size of state space in the model with full information is an order of magnitude less than that of the model with periodic full revelation of information (the FI model states are the cost states of the generators, while the AsI states in the model with periodic full revelation consists of the cost states in the revelation period *and* the history of maintenance decisions since). So if we limit our attention to the recurrent classes the computational demands of the EBE model are *similar* to those from the FI model. If similar results were true for other environments they would imply that we could allow for asymmetric information and use the EBE equilibrium in environments where only symmetric information Markov Perfect models have been used to date. Recall that the EBE equilibrium can be computed in many situations where FI models are computationally infeasible.

On the other hand the sizes of the recurrent class of points in both the AsI model with periodic full revelation of information and for the Markov Perfect symmetric information model are quite large; perhaps not too large for modern computers to handle, but large enough for us to wonder whether agents could keep that much information in their memory without a formal retention and storage process. The size of the recurrent class in the AsI model that uses a finite history of τ to summarize the state is about forty per cent of the size of the recurrent class with full revelation, but there is a question of whether the bounded memory assumption delivers policies with noticeably different implications than the policies from the model with periodic full revelation. To investigate this we ran a one million iteration simulation from the same initial condition for each of our models. Table 3 compares summary statistics from the three AsI models. All summary statistics are virtually identical across the three models (and this was also true for the statistics on policies we consider below).

So 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 restricted information do quite well in approximating unconstrained equilibrium behavior, and they have distinct behavioral advantages. Of course this leaves open the question of how to efficiently construct information sets, a problem whose solution is likely to require institution specific knowledge, but whose implications are likely to be empirically testable.

Finally, since the differences between the AsI models are so small, the

remainder of the paper only presents results from one of them (the model with periodic full revelation).

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

4.5 The Economics of the Alternative Environments.

The output of the algorithm includes a recurrent class of states as well as the associated; strategies, realized costs (both operational and maintenance), profits, and consumer welfare. We begin with the maintenance decisions and use the solution to the social planner problem as our reference point.

The Social Planner Problem. The solution to the social planner problem provides a basis for understanding the logic underlying efficient maintenance decisions for our parameterization. Recall that there is significantly less demand on weekends than on weekdays. Table 4 presents average shut-down 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 $\omega = 1$).

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 AsI								
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

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 (Markov Perfect) solution leaves the two firms with one of two combinations of operating generators over 70% of the time on *each* weekday; about 45% of the time there are two of each type 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 occurring 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 than on weekdays. Apparently the incentives facing an unregulated duopoly with

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

	Dist. ω Prior to Shutdown.					Maint* Freq.
	$\omega = 0$	$\omega = 1$	$\omega = 2$	$\omega = 3$	$\omega = 4$	
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 AsI						
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.

full information lead to maintenance decisions which are very different than those favored by the social planner.

The Duopoly with Asymmetric Information. Perhaps the most striking finding in Table 5 is that there is so much less maintenance in the AsI model than in the FI model. Indeed our “maintenance frequency” summary statistics from the AsI 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 AsI duopoly than the social planner does, and though the firm with the small generators does more maintenance, it does about 25% less than the small firm does in the FI duopoly.

There are several possible reasons for the differences between the AsI and FI withdrawal strategies. Their recurrent classes could differ, or the static incentive to bid additional generators could be convex in the bids of a firm’s competitor inducing less maintenance in the AsI model (we have checked and confirmed this on our recurrent class). Alternatively, the AsI model which allows strategies to depend on past bids could, at least in principle, enable co-ordination (though on imperfect information). However the net result

appears to be that 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.

The fact that the overall maintenance frequency in the AsI equilibrium is similar to the social planner’s maintenance frequency, does not, however imply that the *distribution* of maintenance dates in the AsI equilibrium is optimal. In particular the AsI equilibrium generates more shutdowns for the firm with the small generators 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 operates²⁷. This clearly indicates that the strategies were influenced by the abilities of firms to coordinate their decisions through the use of maintenance decisions as signals.

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

	Duopoly AsI	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

²⁷This was particularly true of the firm with the small generator. In the duopoly with AsI, 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.

Consumer Surplus and Profits. Given the maintenance patterns it is not surprising that the full information equilibrium generates less consumer surplus than the asymmetric information equilibrium. The FI equilibrium generates only 88-89% of the consumer surplus generated by the planner while the AsI equilibrium generates 92-93% of the planner's consumer surplus. The FI model does do better on profits than the AsI model, so the FI model's total surplus is only 6.5% less than that of the social planner, while the asymmetric information equilibrium generates a surplus which is 4% less than the social planner. Clearly then this is a case where society would prefer that firms have less information.

At least two other points are worthy of note. Our exercise 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 AsI equilibria magnify this difference. This suggests that institutions which change the pattern of withholding generators are unlikely to do away with 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%). So 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 equilibrium conditions which, at least in principal, are empirically testable, and an algorithm capable of computing policies which satisfy those conditions for a given set of primitives. Its advantages are twofold. First by choosing alternative information structures we can approximate behavior by agents in complex institutional settings without requiring those agents to have unrealistically excessive information retention and computational abilities. Second the algorithm we use for analyzing the equilibria 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. The hope is that this will enable us to approximate behavior and analyze outcomes in markets which have been difficult to deal with to date. This includes markets with dynamic consumers as well as dynamic producers, and markets where accounting for persistent sources of asymmetric information is crucial to the analysis of ourcomes.

References

- Athey, S., and K. Bagwell, 2008, "Collusion with Persistent Cost Shocks," *Econometrica*, May 2008, 76 (3), 493-540.
- Bajari, P., L. Benkard and J. Levin, 2007, "Estimating Dynamic Models of Imperfect Competition", *Econometrica*, 75, pp.1331-70.
- Benkard, L. (2004), "A Dynamic Analysis of the Market for wide-bodied Commercial Aircraft" *RESTUD*, 71 (3) pp. 581-611.
- Beresteanu A. and P. Ellickson (2006) *The Dynamics of Retail Oligopoly*, mimeo Department of Economics, Duke University.
- Besanko, D., U. Doraszelski, Y. Kryukov and M. Satterthwaite (2010), "Learning by Doing, Organizational Forgetting and Industry Dynamics" *Econometrica*, 78 (2) pp.453-508.
- Borenstein S., J. Bushnell, and F. Wolak (2002) "Measuring Market Inefficiencies in California'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*, 98, pp 114-126.
- Dekel, E., D. Fudenberg and D.K. Levine (2004), "Learning to play Bayesian Games" *Games and Economic Behavior*, 46, 282-303.
- Doraszelski, U. and S. Markovich (2007), "Advertising Dynamics and Competitive Advantage" *Rand Journal of Economics*, 38 (3) pp. 1-36.
- 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 (2010), "Computable Markov-Perfect Industry Dynamics", *RAND Journal of Economics*, pp. 215-43.
- 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.K. Levine (1993) "Self Confirming Equilibrium" *Econometrica*, 61(3), 523-545.
- 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.
- Gowrisankaran, G. and R. Town, (1997) "Dynamic Equilibrium in the Hospital Industry" *Journal of Economics and Management Strategy*, 6 (1) pp. 45-74.
- 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.
- Lee R. and A. Pakes (2009) "Multiple Equilibria and Selection by Learning in an Applied Setting", *Economics Letters*, Volume 104, Issue 1, pp 13-16.
- 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 (1994), "Computing Markov Perfect Nash Equilibrium: Numerical Implications of a Dynamic Differentiated Product Model," *RAND Journal of Economics*, vol. 25, no. 4, pp. 555-589.
- Pakes, A. and P. McGuire (2001), "Stochastic Algorithms, Symmetric Markov Perfect Equilibrium, and the "Curse" of Dimensionality", *Econometrica* pp.1261-81.
- Pakes 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.
- Patrick R. and F. Wolak (1997); "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market" *POWER* working paper PWP-047, University of California Energy Institute ".
- Robbins and Monro (1951); "A Stochastic Approximation Technique", *Annals of Mathematics and Statistics*, 22, pp. 400-07.

Appendix 1: Periodic Revelation of Information.

Claim 1 Periodic Revelation. *If for any initial $s_t \in \mathcal{R}$ there is a $T^* < \infty$ and a random 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 $\#\mathcal{R}$ 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 $(W^*(\cdot|J_i), m^*(J_i), p^e(\cdot|J_i))$ be EBE valuations, strategies, and resulting probability distributions when agents condition both their play and their evaluations on J_i (so they satisfy C1, C2, C3 of section 2). Fix $J_i = J_{i,t}$. what we much show is that

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

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}) ,

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

If this is the case strategies which satisfy the optimality conditions with respect to $\{W^*(\cdot|J_{i,t})\}$ will satisfy the the optimality conditions 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 $Pr(J_{-i,t_0}|h_{i,t_0}) = p^e(J_{-i}|J_{i,t_0})$. What we must show is that if agents use the * policies then the distribution of J_{-i,t_0+1} conditional on h_{i,t_0+1} depends only on J_{i,t_0+1} .

Let a bar over a set of variables indicate its complement in $\cup_i J_{i,t}$ for any t , and

$$\mu_i \equiv J_{i,t_0+1} \cap \bar{J}_{-i,t_0+1} \cap \bar{J}_{i,t_0}, \text{ while } \epsilon \equiv \cap_i J_{i,t+1} \cap_i \bar{J}_{i,t}$$

so that μ_i is the new private, and ϵ is the new public, information in J_{i,t_0+1} . We assume that

$$(A1.) \quad P(\mu_i|h_{i,t}) = P(\mu_i|J_{i,t}, m_{i,t}) \quad \text{and} \quad P(\epsilon|\cup_i h_{i,t}) = P(\epsilon|\cup_i J_{i,t}, \cup_i m_{i,t})$$

so that the distribution of the new private and public information depend only on agents' policies and the information in $\cup_i J_{i,t}$. The fact that (A1) allows the distribution of ϵ to depend on policies generates the possibility of sending signals or revealing information on events that have occurred since all information was revealed. What (A1) rules out is models where the interpretation of those signals depends on information that occurred prior to the period when all states were revealed.

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

$$Pr(J_{-i,t_0+1}|h_{i,t_0+1}) = Pr(\mu_{-i}, \epsilon, J_{-i,t_0}|\mu_i, \epsilon, h_{-i,t_0}) = \frac{Pr(\mu_{-i}, \mu_i, \epsilon, J_{-i,t_0}|h_{i,t_0})}{Pr(\mu_i, \epsilon|h_{i,t_0})}.$$

From (A1) and the * policies, the numerator in this expression can be rewritten as

$$Pr(\mu_{-i}, \mu_i, \epsilon, J_{-i,t_0}|h_{i,t_0}) = Pr(\mu_{-i}, \mu_i, \epsilon, J_{-i,t_0}|\cup_i J_{i,t_0}, \cup_i m^*(J_{i,t_0}))Pr(J_{-i,t_0}|h_{i,t_0}),$$

and from the hypothesis of the inductive argument $Pr(J_{-i,t_0}|h_{i,t_0}) = p^e(J_{-i,t_0}|J_{i,t_0})$. A similar calculation for the denominator concludes the proof. ♠

Claim 2 There exists a restricted EBE if there is periodic revelation of information. ♠

Appendix 2: Algorithmic Details.

We begin with a brief review of the properties of the algorithm, and then move to some notes on how one might usefully amend the algorithm to be more efficient when different primitives are appropriate.

The advantages of using a stochastic algorithm to compute equilibria in full information games relative to standard iterative techniques like those used in Pakes and McGuire (1994) were explored by Pakes and McGuire (2001). These advantages are even larger in asymmetric information games that use the EBE equilibrium conditions. This because those conditions do not require us to form beliefs about player's *types*, and the stochastic algorithm neither computes posterior beliefs nor tests for their consistency with the actual distribution of types.

Pakes and McGuire (2001) noted that, at least formally, their stochastic algorithm does away with all aspects of the curse of dimensionality but that in

computing their test statistic. Accordingly as they increased the dimension of the state space in their examples the computation of the test statistic quickly became 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. There are, however, both choices in implementation, and amendments to the algorithm, that will influence which equilibrium is computed.

One choice is that for the initial evaluations i.e. our W^0 . High initial values are likely to encourage experimentation and lead to an equilibria in which players have explored many alternatives. An alternative way of insuring experimentation is to amend the algorithm as follows. Instead of having agents chose the “greedy” policy at each iteration, that is the policy that maximizes W^k , use choice procedure which has an exogenous probability of chosing each possible action at each early iteration, but let that probability go to zero for all but the greedy policy as the number of iterations grows large. Though both these procedures will insure experimentation, they will also tend to result in longer computational times.

As noted in a particular applied context one may be more interested in directing the algorithm to compute an equilibrium which is consistent with observed data, say by introducing a penalty function which penalizes deviations from the exogenous information available, then in computing an equilibria which insures experimentation. Relatedly note that since our estimates of the $\tilde{W}$ are sample averages, and will be more accurate at a given location the more times we visit that location. 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 from that point.

Both the structure of memory provided and the test given in the text are always available, but that memory structure need not be computationally efficient, and the test need not be the most powerful test. A brief discussion of alternative memory structures and testing procedures follows.

Alternative Memory Structures. It is useful to work with the distribution of the increment in ω between two periods, i.e. defining $\eta_{t+1} \equiv \omega_{t+1} - \omega_t$, we work with

$$\mathcal{P}_\eta = \{ P_\eta(\cdot | m_i, m_{-i}, \omega); (m_i, m_{-i}) \in \mathcal{M}^n, \omega \in \Omega \},$$

where $\mathcal{P}_\eta$ is derived from the family of distributions in equation (2).

We begin with the case where m is observed by the agent's competitors. Then we could hold in memory either estimates of $W(m|J_i)$ or estimates of $W(\eta, m|J_i)$. If the latter we would chose m at iteration k to maximize $\sum_\eta W^k(\eta, m|J_i)p(\eta|m, m_{-i}^{k-1}, \omega)$. The tradeoff here is clear. By holding estimates of $W(\eta, m)$ instead of estimates of $W(m)$ in memory, we increase both memory requirements and the number of summations we need to do at each iteration. However we are likely to decrease the number of iterations needed until convergence, as explicit use of the primitive $p(\eta|\cdot)$ allows us to integrate out the variance induced by η conditional (m, J_i) rather than relying on averaging the simulation draws to do so. The $W(\eta, m|J_i)$ memory structure is particularly easy to use when the probability of η conditional on m_i is independent of m_{-i} (i.e. in capital accumulation games), and we used it in our electric utility example.

When m is unobservable there is an even simpler memory structure that can be used in capital accumulation games. We can then hold in memory estimates of $W(\eta|J_i)$ and chose m at iteration k to maximize $\sum_\eta W^k(\eta|J_i)p(\eta|m, \omega)$ (we can not do this when m is observable because then m is a signal and will have an effect on next period's state that is independent of η). Then the memory requirements may be larger when we hold estimates of $W(m|J_i)$ in memory relative to holding estimates of $W(\eta|J_i)$, and will be if the cardinality of the choice set (of M) is greater than the cardinality of the the support of the family $\mathcal{P}_\eta$. Notice that the model that holds estimates of $W(\eta|J_i)$ in memory is a natural way of dealing with continuous controls (continuous m) whose values are unobserved by competitors, and that we may well have some controls observed and some unobserved in which case hybrids of the structures introduced above would be possible. As for computational burden, the model that holds estimates of $W(\eta|J_i)$ in memory has the advantage that it explicitly integrates out over the uncertainty in η and hence should require less iterations until convergence.

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

Second we used an informal stopping rule, stopping the algorithm when the

norm of the bias in the estimates of $\{W(\cdot)\}$ was sufficiently small. Instead we could have used a formal statistical test of the null hypothesis that there was no bias (i.e. test the null $H^0 : \mathcal{T} = 0$). Notice that if we did proceed in this way we could, by increasing the number of simulation draws, increase the power of any given alternative to one. This suggests that we would want to formalize the tradeoff between size, power, and the number of simulation draws, and explicitly incorporate allowance for imprecision in the computer's calculations. These are tasks we leave to future research.