Most of this lecture is based on “Finite State Dynamic Games with Asymmetric Information: A Framework for Empirical Work”, by Chaim Fershtman and Ariel Pakes.

**Goal:** Provide a framework for applied analysis of dynamic games that allows for *persistent sources of asymmetric information*,

- which generates policies which are both: (i) “relatively” easy for agents to use and (ii) are optimal in some meaningful sense of that word, and
- which is “amenable” to applied research.

“Relatively easy to use” Do not demand as much of agents either; (i) computationally and/or (ii) on information access & retention, as perfect Bayesian Nash conditions.
“Amenable to applied research”. Equilibrium conditions are:
(i) defined in terms of observable magnitudes and hence can be tested (at least in principal),
(ii) its policies can be computed with relative ease (in particular they should be computable without the curse of dimensionality, as this limits what we can do in applied I.O.).

Some Applications.

• Analysis of markets with “dynamic demand” (e.g. durable or experience goods). Reasonable to assume;
  (i) consumers and producers have different information sets (producers know more about costs and often gather demand side information that is treated as proprietary),
  (ii) neither have information on all “relevant state variables” (e.g. entire joint distribution of past purchases and consumer characteristics), and even if they did one might wonder whether we can get a better approximation to behavior than a Perfect Bayes-Nash solution.

• Dynamic games between producers with persistent sources of asymmetric cost info. Current e.g.: maintenance decisions of electric utilities.

  Use other examples to illustrate specific points below.

“Testability” rules out conditions formulated in terms of Bayesian posteriors. Use conditions based on observed outcomes (similar to Fudenberg and Levine, 1993). Implies

• Limitation. Only likely to provide an adequate approximation when the returns from the possible actions can be
learned from pervious play (does provide a learning rule with a direct interpretation).

- **Limitation.** Since the returns differ with the state, and the states evolve over time, this essentially limits us to games with finite state spaces (we discuss how we might justify that below).

The finiteness condition has been a concern for theoreticians in different ways (come back to this below), but has been less of a concern for applied projects.

**Structure of game.** Realized profits are

$$\pi_i(\omega_{i,t}, \omega_{-i,t}, m_{i,t}, m_{-i,t}, x_{i,t}, d_t),$$

where

- $\omega$ are “payoff relevant” states. Firm knows its own state but *may not know its competitors’* states,

- $m$ is a “control” with finite support; a firm knows its own value but *may or may not know* its competitors (if it knows its competitors’ $m$ but not $\omega$, the $m$ becomes a signal on the competitor’s unknown $\omega$ state)

- $x$ is a “control” with possibly continuous support; known only to the firm (not to its competitors)

- $d$ are publically observed states with values that are common to all firms.
Simple Electric Utility Eg.

*Two firms*: each has a vector of generators.

*Firm’s decisions*: bid or not each generator.

*ISO*: sum bid functions, intersect with demand (varies by day of the week), pay a uniform price to accepted electricity.

- $\omega \in \Omega$. Cost of producing electricity on each firm’s generators. Cost increases stochastically with use, but reverts to a starting value if the firm goes down for maintenance.

- $m \in M$. Vector of zero-one variables: indicates which generators the firm bids this period. If it does bid, it bids a fixed schedule, and if it does not bid it goes down for maintenance.

- $d$ is demand on that day, $f$ is maintenance cost (“investment”), $p = p(m_i, m_{-i}, d)$ is price, $q = q(m_i, m_{-i}, d)$ is allocated quantity, so realized profits are

$$
\pi_{i,t} = \sum_r p_t q_{i,r,t} - \sum_r c_i(\omega_{i,r,t}, q_{i,r,t}) - f_i \sum_r [1 - m_{i,r,t}]
\equiv \pi_i(\omega, m_i, m_{-i}, d)
$$

$m_{i,r,t} = 0 \Rightarrow \omega_{i,r,t+1} = \omega_{i,r}$ (\omega=restart state).

$m_{i,r,t} = 1 \Rightarrow \omega_{i,r,t+1} = \omega_{i,r,t} - \eta_{i,r,t}$ with $P(\eta) > 0$ for $\eta \in \{0, 1, 2, 3\}$.

**Note** $m$ is the only signal sent in each period. $m_{-i,t-1}$ is a signal on $\omega_{-i,t-1}$ which is unobserved to $i$ and is a determinant of $m_{-i,t}$ (and so $\pi_{i,t}$).
Cases.

• Example. $\pi_i = \pi_i(\omega_i, m_i, m_{-i}, d)$. Only $m_i$ effects the evolution of $\omega_i$. It is also the only signal sent in each period. $\omega_{-i}$ is unobserved and effects $m_{-i}$ (and so $\pi_i$).

• Capital accumulation games.
  $\pi_i = \pi_i(\omega_i, m_i, m_{-i}, x_i, d)$ with $(\omega_{-i}, m_{-i})$ possibly unknown. Only own controls, $(m_i, x_i)$, effect evolution of $\omega_i$. Focus of presentation.

• General. $m_{-i}$ can also effect the distribution of $\omega_i$. Needed for some auction and learning by doing games. Appendix of paper.

State of the game. $s_{i,t} = (J_{1,t}, \ldots J_{n_i,t}) \in \mathcal{S}$, where $J_{i,t}$ is the information set (and hence determines policies) of firm $i$ in period $t$. 

$$J_{i,t} = (\xi_{i,t}, z_{i,t}) \in (\Omega(\xi), \Omega(z))$$

where

• $z_{i,t}$ represents private information (known only to $i$). Example: $\omega_{i,t} = z_{i,t}$.

• and $\xi_t$ is public information (shared by all). Example $\xi_t = \{m_{1,\tau}, m_{2,\tau}, d_\tau\}_{\tau \leq t}$.

Note. Without further restrictions all past history is informationally relevant. Example: A’s decision depends on whether it thinks $B$ will bid in. This depends on the last time B went down. But the last time B went down depends on the time A went down before that.... Infinite regress.
Assume. Every $L$ periods a regulator inspects machines and files a public report. Can show that then there are equilibrium strategies that only depend on the report and the bids and demand since. If we restrict to those strategies we get a finite state space. Consider other ways of generating finiteness below.

Complexity of Restricted PBE. $i$’s policies at $t = 1$ change $-i$’s posteriors at $t = 2$, so to calculate equilibrium policies we need continuation values. Continuation values require us to solve jointly for: (i) equilibrium policies

$$m_{i,l}(\omega_{i,0}, \omega_{-i,0}, \{\eta_{i,\tau}, m_{i,\tau}, m_{-i,\tau}\}_{\tau=1}^{l}),$$

and (ii) the posterior distributions they imply

$$Pr\{m_{-i,l}(\cdot)|(\omega_{i,0}, \omega_{-i,0}, \{\eta_{i,\tau}, m_{i,\tau}, m_{-i,\tau}\}_{\tau=1}^{l})\},$$

for every sample path

$$(\omega_{i,0}, \omega_{-i,0}, \{\eta_{i,\tau}, m_{i,\tau}, m_{-i,\tau}\}_{\tau=1}^{l}),$$

and every $l \leq L$.

Alternative equilibrium concept: “Applied Markov Equilibrium”. For simplicity assume $\eta_{i,t} = \omega_{i,t} - \omega_{i,t-1}$, i.e. $\eta$ is the “innovation” in the payoff relevant state, and let the distribution of $\eta$ depend only on $(x_{i,t}, \omega_{i,t-1})$ (so this is a capital accumulation game).
Then an equilibrium is a triple

- A subset of industry states \( \mathcal{R} \subset \mathcal{S} \);
- Strategies \((x^*(J_i), m^*(J_i))\) for every \(J_i\) which is a component of any \(s \in \mathcal{S}\);
- Expected discounted values of net cash flows conditional on realizations of \(\eta\) and a choice for \(m\), say \(\{W(\eta, m|J_i)\}_{\eta, m}\), \((\forall s \in \mathcal{S}, J_i \in s\) )

such that

**C1:** \(\mathcal{R}\) is a recurrent class of the Markov process for \(\{s_t\}\) generated by the outcomes from the optimal policies (from \(\{(x^*, m^*)\}\)). \((\Rightarrow \text{with probability one, any subgame starting from an } s \in \mathcal{R} \text{ will generate sample paths that are within } \mathcal{R} \text{ forever.})\)

**C2:** Optimality of strategies on \(\mathcal{R}\). Recall \(W(\eta, m|J_i)\) is the expected discounted value if we chose \(m\) and the stochastic outcome is \(\eta = \omega_{i,t} - \omega_{i,t-1}\). For strategies to be optimal given these values we need \((x^*(J_i), m^*(J_i))\) to solve

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

\(\forall J_i \subset s \in \mathcal{R}\).

**C3:** Consistency of values (or \(W(\cdot)\)) on \(\mathcal{R}\). The values \((W(\cdot))\) assigned to outcomes that are observed repeatedly are consistent the values generated by equilibrium play.
Formally let

- \( \eta \in \eta(x^*(J_i), m^*(J_i), \omega(J_i)) \) be the set of \( \eta \) values that have positive probability given equilibrium policies, and
- \( p^E(\cdot|\cdot) = \) the (limiting) **empirical distribution** of outcomes.

Then \( \forall \eta \in \eta(x^*(J_i), m^*(J_i), \omega(J_i)) \)

\[
W(\eta, m^*(J_i)|J_i) = \pi^E(J_i) + \beta \sum_{J'_i} \left\{ \sum_{\tilde{\eta}} W(\tilde{\eta}, m^*(J'_i)|J'_i)p(\tilde{\eta}|x^*(J'_i), \omega(J'_i)) \right\} p^E(J'_i|J_i, \eta),
\]

where

\[
\pi^E(J_i) \equiv \sum_{J_i} \pi_i(\omega_i, m^*_i, x^*_i, \omega_{-i}, m^*(J_{-i}), d_i)p^E(J_{-i}|J_i),
\]

and

\[
\left\{ \begin{array}{c}
p^E(J'_i|J_i, \eta) \equiv \frac{p^E(J'_i, \eta|J_i)}{p^E(\eta|J_i)} \quad \text{for } J'_i, \\
p^E(J_{-i}|J_i) \equiv \frac{p^E(J_{-i}, J_i)}{p^E(J_i)} \quad \text{for } J_{-i}
\end{array} \right. \]
Notes on the equilibrium conditions.

**Learning.** C3 states that if $Pr(\eta|m^*(J_i)) > 0$, $W(\eta, m^*|J_i)$ must equal the average of the values that would actually be generated by equilibrium play. ⇒ if the game is played repeatedly one could learn what $W(\cdot)$ is from past outcomes.

**Additional Constraints.** Depending on the institutional structure we are trying to approximate it may make sense to impose additional constraints. In our simple example $m_{-i}$ is observed by agent $i$, so when we compute our equilibrium we add the condition that, for $m \neq m^*$

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

**Boundaries of $\mathcal{R}$**. An $s \in \mathcal{R}$ for which there are $\tilde{\eta} \notin \eta(x^*(J_i), m^*(J_i), \omega(J_i))$, are boundary points. Equilibrium behavior: $s \in \mathcal{R} \Rightarrow s' \in \mathcal{R}$ w.p.1, but at boundaries there are feasible strategies from $s \rightarrow s' \notin \mathcal{R}$.

**Testing.** Assume the empiricist knows the union of the information sets of all players $\forall t$ (this is the best we could expect). Still, at boundary points there are some $(\tilde{\eta}, \tilde{m})$ that are not observed repeatedly, so we could never construct a consistent test of whether their $W(\tilde{\eta}, \tilde{m}|J_i)$ is rationalized by subsequent behavior. The only conditions on these $W(\cdot)$ are inequalities which insure optimality of policy.
**Multiplicity of Equilibria.** Enhanced possibility. But (generically) unique equilibrium strategies on a given $R$ given $W$’s (which are estimable).

**Finite State Space Condition.**

Conditions on primitives can insure that “payoff relevant” states (the $\omega$) take values only on a finite space. Harder for “informationally” relevant states. How to insure it?

**Institutions which generate it:**
(i) Periodic full revelation of information,
(ii) Functional forms (finite dimensional sufficient statistics for unknowns),
(iii) Agents only have access to a finite history (e.g. e-bay).

**Bounded cognitive abilities** limit the game to a finite state space (e.g. bounded memory).

*Investigate whether this leads to policies which approximate the equilibria to an unconstrained game in our example.*

To go this route we have to decide what is kept in memory (empirically testable?).

The computational burden is (essentially) the product of three factors,

- the number of points evaluated at each iteration;
- the time per point evaluated;
- the number of iterations.

Number of Points.
Since each of the $\pi$ active firms can only be at $K$ distinct states, the number of points we need to evaluate at each iteration, or

$$\#S \leq K^\pi.$$

Exchangeability (often referred to as symmetry), of the value and the policy functions in the state variables of a firm’s competitors implies that we do not need to differentiate between two vectors of competitors that are permutations of one another. Pakes (1993) shows that an upper bound for $\#S$ is given by the combinatoric

$$\binom{K+\pi-1}{\pi} \ll K^\pi.$$

but for $\pi$ large enough this bound is tight.

Burden per Point.
Determined by
• the cost of calculating the expected value of future states conditional on outcomes (of obtaining the $w^j(\cdot; i, s)$ from the information in memory).

• The cost of obtaining the optimal polices and the new value function given $w^j(\cdot; i, s)$.

Take the simplest model and recall that

$$V^j(i, s) = \max_{\chi \in \{0, 1\}} \{[1-\chi]\phi + \chi\{\pi(i, s) - \sup_{x \geq 0}[-cx + \beta \sum_{\nu} w^{j-1}(\nu; i, s)p(\nu|x)]\}\},$$

where

$$w^j(\nu; i, s) \equiv \sum_{(\hat{s}', \zeta)} V^{j-1}(i+\nu-\zeta, \hat{s}'+e(i+\nu-\zeta)|w)q^{j-1}[\hat{s}'|i, s, \zeta] \mu(\zeta),$$

and

$$q^{j-1}[\hat{s}' = s^*_i|i, s, \zeta] \equiv \Pr\{\hat{s}' = s^*_i|i, s, \zeta, \text{policies at iteration "j-1" } \}.$$

Easiest to think of this as individual firms instead of a measure, i.e as a state being $(i_j, \bar{i})$. Assume that there is positive probability on each of $\kappa$ points for each of the $m-1$ active competitors of a given firm. Then we need to sum over $\kappa^m$ possible future states and there are $\kappa \times m$ values of $w^j(\cdot)$ needed at that $s$. Average $m$ should increase in $\bar{n}$, and $\kappa$ should be determined by the nature of the state space per firm (it typically goes up exponentially in the number of state variables per firm).
Conclusion

Little known about the relationship between the number of iterations and the number of state variables of the problem; but doesn’t seem to be too bad. Still it is clear that the computational burden of the model grows quickly in both the number of firms ever simultaneously active (it grows geometrically in this dimension) and the number of state variables per firm (it grows exponentially in this dimension). This is the problem known as “The Curse of Dimensionality” in the computational literature.

Approximation Techniques.

- Pointwise algorithm has been used both as a tool for substantive problems and as a teaching device.
- Still many applied problems need more powerful computational tools.
- Three available. Each has their problems, but they compute equilibria with much less of a computational burden than the standard algorithm.
  - Continuous time algorithm. Doraszelski and Judd (forthcoming).
  - Deterministic approximation techniques. Judd (book)
- There are also improved computer programs for finding fixed points (MPEC), but they do not seem to be able to handle fixed points of the size we often need.
I do not know of a publicly available versions of these algorithms, though as we shall they are often not harder, and can be easier (especially the stochastic algorithm), then the “brute force” algorithm we initially discussed.

I will outline what happen in the deterministic approximation and continuous time versions, and will apply the stochastic algorithm to the dynamic game with asymmetric info.

Continuous Time.

The first approach we describe, due to Doraszelski and Judd, is designed to ease the burden of computing the expectation over successor states, and hence decreases the computation time needed to update values and policies at a particular state (it makes no attempt to alleviate the burden imposed by the number of states). The set up is a continuous-time model in which at any particular instant only one firm experiences a change in its state. As a result if each firm’s transition can go to one of \( K \) states, and there are \( n \) firms, we only need to sum over \( (K - 1) \times n \) states to compute continuation values. This is in contrast to the \( K^n \) possible future states that we need to sum over in the computational algorithm described earlier.\(^1\) This implies that the discrete- and continuous-time models have different implications. As a result it may (but need not) be the case that one of the models provides a better approximation to

\(^1\)Note that if we are willing to explicitly restrict players to move one at a time we could obtain similar gains from a discrete-time model in which decisions are made sequentially as a result of a random selection mechanism and outcomes are realized before the next decision is made. From a computational point of view the deterministic order prevents us from using anonymity to reduce the size of the state space. However use of a random order of moves would preserves anonymity.
behavior in a particular setting than the other. For example in the continuous time model starting at $t_1$ which firm changes its state is a probabilistic function of all firm’s investments. However if firm 1 changes its state at $t_2$ none of the investments of the other firms in the interim affect future sample paths. In the discrete time model the investments of other firms are totally captured in the current state, so investment prior to $t_1$ do not matter conditional on the state, but investments in the interval do.

**Deterministic Approximations (“Curve Fitting”).**

Judd’s (2000) book provides a thorough introduction to such techniques. The techniques begin by specifying a set of functions considered rich enough to contain an element which provides a good approximation to the value function (e.g. polynomials of order $d$). Given some initial value functions for a small subset of the points in $\mathcal{S}$, they then find a member of the set of functions that approximates the value functions at the small set of points (in a polynomial they would find a number of points that is large and “varied” enough to enable them to set the polynomial coefficients). They then do the iteration on the small set of points using the approximating function to predict the value function at other points as needed. They then continue until convergence.

An example is given at the end of the PM(1994) paper. The paper contains a fairly detailed description of how to construct an algorithm based on these techniques for our problem. This shows that the tools that can be brought to our problem are quite powerful. However we were not terribly successful when
we used them on problems the size of the problem computed in PM(1994). This is possibly because of our use of polynomials and the fact that the value function is likely not to be “smooth” in the values of the competitors states. The techniques are likely to be more successful when \( \pi \) is larger, where one would expect lack of smoothness to be a smaller problem. Indeed Liu’s thesis(1998) uses a combination of the polynomial approximation technique and simulation, and successfully computes equilibria for a 35 firm industry.

**Computational Algorithm for AME.**

We will introduce a reinforcement learning algorithm for this asymmetric info game (might be viewed “approximation” to a learning process). A similar algorithm can be used for a game of full information (see Pakes and McGuire (2000, *Econometrica*). The algorithm has neither curse of dimensionality in computation or in testing. Other methods or reducing the computational burden are also available, and if I have time I will go over them.

**Iterations.** Defined by:

- A location, say \( L^k = (J^k_1, \ldots J^k_{n(k)}) \in S \): is the information sets of agents active.
- Objects in memory (i.e. \( M^k \)):
  (i) perceived evaluations, \( W^k \),
  (ii) No. of visits to each point, \( h^k \).

So algorithm must update \( (L^k, W^k, h^k) \).
Update Location.
- Calculate policies for all agents active at $L^k$ to maximize the agents' values conditional on the evaluations in memory ($W^k$).
- Take random draws on outcomes conditional on those policies and $L^k$ (in our example on $\eta$ conditional on $m$) and use them to form $\{J_{i}^{k+1}\}_i$.

Update $W^k$. Define the agent’s ex post perception of what its value would have been had it chosen $m$ and drawn $\eta_i = \eta$ to be

$$V^{k+1}(J_i^k, \eta, m) = \pi(\omega_i^k, \omega_{-i}^k, m, m_{-i}^k, x^k, d^k) +$$

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

where $J_{i}^{k+1}(\eta, m)$ is what the $k+1$ information would have been given: $(\eta, m)$ and competitors actual play.

Treat $V^{k+1}(J_i^k, \eta)$ as a random draw from the possible realizations of $W(\eta, m | J_i^k)$, and update $W^k$ as we update averages

$$W^{k+1}(\eta, m | J_i^k) - W^k(\eta, m | J_i^k)$$

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

Note. If we are in equilibrium we tend to stay their. I.e. if $^*$ indicates equilibrium values and policies, then

$$\forall \eta \in \eta(x^*(J_i), m^*(J_i), \omega(J_i)), \text{ and, } \forall J_i \subset s \in \mathcal{R}$$

$$E[V^*(J_i^k, \eta, m) | W^*, m^*, x^*] = W^*(\eta, m | J_i^k),$$

so if we are in equilibrium we tend to stay their.
Computational properties

• No proof of convergence, though we can test if we have converged (see below).

• Algorithm is asynchronous. Eventually wanders into an \( \mathcal{R} \) and stays their. \( \#\mathcal{R} \) not necessarily related to \( \#\mathcal{S} \). Kills one source of curse of dimensionality. In practical I.O. problems computational burden tends to go up linearly (rather than exponentially or geometrically) in number of state variables.

• No curse of dimensionality in calculating continuation values because of stochastic integration (all we do is form an average).

Testing For an Equilibrium.

Any fixed \( W \), say \( \tilde{W} \), generates policies which define a finite state Markov process for \( \{ s_t \} \). Gather the transition probabilities into the Markov matrix, \( Q(s', s|\tilde{W}) \). Want to test if process satisfy our equilibrium conditions.

For Test  Need 
(i) a candidate for \( \mathcal{R} \), and checks for 
(ii) optimality of policies and 
(iii) consistency of \( W \).

Candidate for \( \mathcal{R}(\tilde{W}) \). Start at any \( s^0 \) and use \( Q(\cdot, \cdot|\tilde{W}) \) to simulate a sample path \( \{ s^j \}^{J_1+J_2}_{j=1} \). Let \( \mathcal{R}(J_1, J_2, \cdot) \) be the set of
states visited at least once between $j = J_1$ and $j = J_2$.

$$(J_1, J_2) \to (\infty, \infty) \Rightarrow \mathcal{R}(J_1, J_2, \cdot) \to \tilde{\mathcal{R}}$$

a recurrent class of $Q(\cdot, \cdot|\tilde{W})$ (C1 satisfied).

**C2: optimality of policies.** Satisfied by construction, since we use the policies generated by $\tilde{W}$ to form $Q(\cdot, \cdot|\tilde{W})$.

**C3: consistency of $\tilde{W}$ with outcomes.** Does

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

$(\forall \eta \in \eta(x^*(J_i), m^*(J_i), \omega(J_i)), \text{ and } J_i \in s \in \mathcal{R}.)$?

*Direct summation.* Computationally burdensome; indeed brings the curse of dimensionality back in.

*Alternative which circumvents computational problem.* Check for consistency of simulated sample paths with evaluations.

- Start at $s_0 \in \mathcal{R}$ and forward simulate. At each $J_i$, keep track of the average and the sample variance of the simulated values, say

$$\left( \hat{\mu}(W(J_i, \eta_i)), \hat{\sigma}^2(W(J_i, \eta_i)) \right) \equiv \left( \hat{\mu}(W_i), \hat{\sigma}^2(W_i) \right).$$

- Note that if we let $E(\cdot)$ take expectations over the simulated random draws and define
\[ \hat{\Upsilon} \equiv E \left( \frac{\hat{\mu}(W_l) - \bar{W}_l}{\bar{W}_l} \right)^2 \]

\[ = E \left( \frac{\hat{\mu}(W_l) - E[\hat{\mu}(W_l)]}{\bar{W}_l} \right)^2 + \left( E[\hat{\mu}(W_l)] - \bar{W}_l \right)^2. \]

\[ = \%Var(\hat{\mu}(W_l)) + \%Bias^2(\hat{\mu}(W_l)). \]

- If \( h_l \) is the empirical measure of visits to \( l \), then as the number of simulation draws grows

\[ \sum_l h_l \left( \frac{\hat{s}_l^2(W_l)}{\bar{W}_l^2} \right) - \sum_l h_l \left( \frac{\hat{\mu}(W_l) - E[\hat{\mu}(W_l)]}{\bar{W}_l} \right)^2 \to_{a.s.} 0, \]

\[ \Rightarrow \hat{\Upsilon} - \sum_l h_l \left( \frac{\hat{s}_l^2(W_l)}{\bar{W}_l^2} \right) \to_{a.s.} \sum h_l \left( \frac{E[\hat{\mu}(W_l)] - \bar{W}_l}{\bar{W}_l} \right)^2, \]

an \( L^2(\mathcal{P}_R) \) norm in the percentage bias (\( \mathcal{P}_R \) is the invariant measure associated with \( \bar{W} \)).
Maintenance in an Electricity Market

Computational Question 1.  Do reasonable bounds on memory lead to policies which approximate unbounded policies? Compare:
(i) AI with full revelation every 6 periods
(ii) AI with 6 periods of memory
(iii) AI remembering last time each generator did maintenance (≤ 6 periods).
Note.  Info(iii) ⊂ Info(ii) ⊂ Info (i) (and provides a solution with no bound on memory).

Computational Question 2.  How do the computational burden of the Full Information and AI games compare? Use reinforcement learning algorithms for all calculations.

Electric Utility Question.  To what extent do maintenance decisions accentuate price hikes in periods of high demand, and does asymmetric information ameliorate or intensify this problem? Compare:
(i) Asymmetric Info (AI) Equilibrium to
(ii) Full Info (FI) Equilibrium to
(iii) Social planner solution.
Model Details.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Firm B</th>
<th>Firm S</th>
</tr>
</thead>
<tbody>
<tr>
<td># Generators</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Range of $\omega$</td>
<td>0-4</td>
<td>0-4</td>
</tr>
<tr>
<td>MC Constant*</td>
<td>(20,60,80,100)</td>
<td>(50,100,150,200)</td>
</tr>
<tr>
<td>Max Mgwt at MC</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>Maint. Cost</td>
<td>15,000</td>
<td>6,000</td>
</tr>
</tbody>
</table>

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

Firm S: small (gas fired) generators with high MC but low start up costs.
Firm B: large (coal fired) generators lower MC and higher start up costs.
MC is constant until a “capacity constraint” and increases thereafter.
Constant, small, elasticity of demand.

Computational Details  Initial conditions. Set high $\Rightarrow$ experimentation.
\[
\pi^E_{i,k=0}(m_i, J_i) = \pi_i(m_i, m_{-i} = 0, d, \omega_i), \text{ and }
\]
\[
W^k=0(\eta_i, m_i|J_i) = \frac{\pi_i(m_i, m_{-i}=0, d, \omega_i + \eta_i(m_i))}{1-\beta}.
\]
Convergence. Criteria: $L^2(\mathcal{P}(\mathcal{R})) \geq .995$. Converges between 100 and 200 million iterations.
**Question 1; Investigate Bounds on Memory.**

<table>
<thead>
<tr>
<th></th>
<th>Finite History of $\tau$</th>
<th>Periodic $m$</th>
<th>Revelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary Statistics.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumer Surplus</td>
<td>2.05 e+07</td>
<td>2.05 e+07</td>
<td>2.05 e+07</td>
</tr>
<tr>
<td>Profit B</td>
<td>2.46 e+06</td>
<td>2.46 e+06</td>
<td>2.45 e+06</td>
</tr>
<tr>
<td>Profit S</td>
<td>2.32 e+06</td>
<td>2.32 e+06</td>
<td>2.33 e+06</td>
</tr>
<tr>
<td>Maintenance Cost B</td>
<td>2.28 e+05</td>
<td>2.28 e+05</td>
<td>2.28 e+05</td>
</tr>
<tr>
<td>Maintenance Cost S</td>
<td>1.66 e+05</td>
<td>1.66 e+05</td>
<td>1.65 e+05</td>
</tr>
<tr>
<td>Production Cost B</td>
<td>2.40 e+06</td>
<td>2.40 e+06</td>
<td>2.39 e+06</td>
</tr>
<tr>
<td>Production Cost S</td>
<td>2.82 e+06</td>
<td>2.83 e+06</td>
<td>2.83 e+06</td>
</tr>
</tbody>
</table>

**Conclude:** computed equilibrium does not differ in meaningful ways when we impose the memory bounds.
Question 2: Computational Comparisons.

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

One Conclusion:  *Compute time not problematic.*
Largely determined by burden of finding points in memory (storage: public information by tree structure, and private information with a hash table conditional on public info).

Second Conclusion.  *Memory requirements may well be problematic.*
AI does *about the same* as FI, and finite history $\tau$ easily wins. Still *all are large* and that is problematic in at least two ways; (i) it will make it hard to analyze more complex games, and (ii) it makes it hard to believe that agents actually condition on this much information.

Question: are their reasonable ways to reduce memory requirements that would both decrease computational burden and also allow us to better approximate behavior?
Question 3: Social Planner vs. AI vs. FI.

<table>
<thead>
<tr>
<th></th>
<th>SP</th>
<th>AI</th>
<th>FI</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Maint when $\omega = {3, 4}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm B</td>
<td>.93</td>
<td>.97</td>
<td>.41</td>
</tr>
<tr>
<td>Firm S</td>
<td>.84</td>
<td>.57</td>
<td>.42</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ave. Operating Generators</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Firm B Wkday</td>
<td>2.37</td>
<td>2.27</td>
<td>1.83</td>
</tr>
<tr>
<td>Firm B Wkend</td>
<td>1.56</td>
<td>2.16</td>
<td>1.91</td>
</tr>
<tr>
<td>Firm S Wkday</td>
<td>3.05</td>
<td>2.43</td>
<td>2.30</td>
</tr>
<tr>
<td>Firm S Wkend</td>
<td>2.43</td>
<td>3.11</td>
<td>2.53</td>
</tr>
</tbody>
</table>

**SP.** Maintenance primarily on weekends. Almost never does maintenance in low cost states.

**FI.** 60% of the time does maintenance at a low cost state. Maintenance about same on weekends and weekdays.

**AI.** Maintenance is done at high cost states (esp. for firm B; uncertainty increases static incentive to bid). However: (i) sometimes shuts down *wrong* generators, (ii) maintenance is done more on weekdays than on weekends.

<table>
<thead>
<tr>
<th></th>
<th>AI</th>
<th>FI</th>
<th>SP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cons. Surplus ($\times 10^{-6}$)</td>
<td>20.51</td>
<td>19.70</td>
<td>22.21</td>
</tr>
<tr>
<td>Prod. Surplus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Firm B ($\times 10^{-6}$)</td>
<td>2.45</td>
<td>2.11</td>
<td>1.99</td>
</tr>
<tr>
<td>Firm S ($\times 10^{-6}$)</td>
<td>2.33</td>
<td>2.83</td>
<td>2.13</td>
</tr>
<tr>
<td>Firms B + S ($\times 10^{-6}$)</td>
<td>4.78</td>
<td>4.95</td>
<td>4.12</td>
</tr>
<tr>
<td>Total Surplus</td>
<td>25.29</td>
<td>24.65</td>
<td>26.34</td>
</tr>
<tr>
<td>Prices</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekend</td>
<td>145.77</td>
<td>170.42</td>
<td>152.51</td>
</tr>
<tr>
<td>Weekday</td>
<td>1205.76</td>
<td>1292.83</td>
<td>990.46</td>
</tr>
<tr>
<td>% of Output Produce by Firm with Larger Generators.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekend</td>
<td>.47</td>
<td>.48</td>
<td>.46</td>
</tr>
<tr>
<td>Weekday</td>
<td>.50</td>
<td>.43</td>
<td>.46</td>
</tr>
</tbody>
</table>

- Not surprisingly the AI equilibrium has more CS and less profits than FI. It also does better in total surplus, by $\approx %4$.
- Even the SP has large price swings between weekdays and weekends, perhaps restrictions on withdrawal of generators are not enough to dampen swings; maybe need smoothed demand.