NBER WORKING PAPER SERIES

METHODOLOGICAL ISSUES IN ANALYZING MARKET DYNAMICS

Ariel Pakes

Working Paper 21999
http://www.nber.org/papers/w21999

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2016

This paper is adapted from three lectures given in 2014: (i) The Cowles Lecture given at the North American Econometric Society, (in Minneapolis, June); (ii) A Keynote Address at the International Society for Dynamic Games (in Amsterdam, July), and (iii) A Keynote Address at the Latin American Econometric Society Meetings (in Sao Paulo, November). I have benefited from the comments of several participants in those conferences and from sets of detailed and very useful comments from Michael Dickstein, Liran Einav, Chaim Fershtman, Myrto Kalouptsidi. and Florian Wagener. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Methodological Issues in Analyzing Market Dynamics
Ariel Pakes
NBER Working Paper No. 21999
February 2016
JEL No. L13

ABSTRACT

This paper investigates progress in the development of models capable of empirically analyzing the evolution of industries. It starts with a parallel between the development of empirical frameworks for static and dynamic analysis of industries: both adapted their frameworks from models taken from economic theory. The dynamic framework has had its successes: it led to developments that have enabled us to control for dynamic phenomena in static empirical models and to useful computational theory. However when important characteristics of industries were integrated into that framework it generated complexities which both hindered empirical work on dynamics per se, and made it unrealistic as a model of agent behavior. This paper suggests a simpler alternative paradigm, one which need not maintain all the traditional theoretical restrictions, but does maintain the core theoretical idea of optimizing subject to an information set. It then discusses estimation, computation, and an example within that paradigm.

Ariel Pakes
Department of Economics
Harvard University
Littauer Room 117
Cambridge, MA 02138
and NBER
apakes@fas.harvard.edu

1 Introduction.

It will be helpful if I start out with background on some recent methodological developments in empirical Industrial Organization, concentrating on those either I have been more closely associated with. I start with an overview of what we have been trying to do and then move on to how far we have gotten. This will bring us naturally to the analysis of market dynamics; the main topic of the paper.

Broadly speaking the goal has been to develop and apply tools that enable us to better analyze market outcomes. The common thread in the recent developments is a focus on incorporating the institutional background into our empirical models that is needed to make sense of the data used in analyzing the issues of interest. These are typically the causes of historical events, or the likely responses to environmental and policy changes. In large part this was a response to prior developments in Industrial Organization theory which used simplified structures to illustrate how different phenomena could occur. The empirical literature was trying to use data and institutional knowledge to narrow the set of possible responses to environmental or policy changes (or the interpretations of past responses to such changes). The field was moving from a description of responses that could occur, to those that were "likely" to occur given what the data could tell us about appropriate functional forms, behavioral assumptions, and environmental conditions.

In pretty much every setting this required incorporating

- heterogeneity of various forms into our empirical models,

and, when analyzing market responses

- using equilibrium conditions to solve for variables that firms could change in response to the environmental change of interest.

The difficulties encountered in incorporating sufficient heterogeneity and/or using equilibrium conditions differed between what was generally labeled as "static" and "dynamic" models. For clarity I will use the textbook distinction between these two: (i) static models solve for profits conditional on state variables, and (ii) dynamics analyzes the evolution of those state variables (and through that the evolution of market structure). By state variables here I mean:

the characteristics of the products marketed, the determinants of costs, the distribution of consumer characteristics, the ownership structure, and any regulatory or other rules the agents must abide by. I begin with a brief review of the approach we have taken with static models, as that will make it easier to understand how the dynamic literature evolved.

Static Models. The empirical methodology for the static analysis typically relied on earlier work by our game theory colleagues for the analytic frameworks we used. The assumptions we took from our theory colleagues included the following:

- Each agent's actions affect all agents' payoffs, and
- At the "equilibrium" or "rest point"
  (i) agents have "consistent" perceptions[1], and
  (ii) each agent does the best they can conditional on their perceptions of competitors' and nature's behavior.

Our contribution was the development of an ability to incorporate heterogeneity into the analysis and then adapt the framework to the richness of different real world institutions. This was greatly facilitated by progress in our computational abilities, and the related increased availability of data and econometric methodology. Of particular importance were econometric developments which enabled the use of semi-parametric (see Powell, 1994) and simulation (see McFadden, 1989, and Pakes and Pollard, 1989) techniques. The increased data, computing power and econometric techniques enabled the framework to be applied to a variety of industries using much weaker assumptions than had been used in the theory literature.

Indeed I would make the claim that the tools developed for the analysis of market allocations conditional on the state "variables" of the problem have passed a "market test" for success in an abundance of situations. I come to the conclusion for three reasons.

- First these tools have been incorporated into applied work in virtually all of economics that deals with market allocation when productivity and/or demand is part of the analysis.

[1] Though the form in which the consistency condition was required to hold differed across applications.

- The tools are now used by public agencies, consultancies and to some extent by firms.
- The tools do surprisingly well, both in fit and in providing a deeper understanding of empirical phenomena.

For examples of the last point I note that empirical analysis of equilibrium pricing equations in retail markets that followed Berry, Levinsohn and Pakes (1995) both; (i) typically fit exceptionally well for a behavioral equation and (ii) generated markups which were in accord with other sources of information on markups. Similarly the productivity analysis that followed Olley and Pakes (1996) was able to separate out and analyze changes in aggregate productivity attributable to: (i) increases in productivity at the firm level and (ii) increases resulting from a reallocating output among differentially productive firms.

I do not want to leave the impression that there is nothing left to be done in the analysis of equilibrium conditional on state variables. There have been several recent advances which have enhanced our ability to use static analysis to analyze important problems. This includes; (i) the explicit incorporation of adverse selection and moral hazard into the analysis of insurance and capital markets (for e.g. Einav, Jenkins, and Levin, 2012), (ii) the analysis of upstream contracts in vertical markets characterized by bargaining (see Crawford and Yurukoglu, 2013). and (iii) the explicit incorporation of fixed (and other non-convex) costs into the analysis of when a good will be marketed (see Pakes, Porter, Ho and Ishii, 2015).

Dynamic Models. Empirical work on dynamic models proceeded in a similar way to the way we proceeded in static analysis; we took the analytic framework from our theory colleagues and tried to incorporate the institutions that seemed necessary to analyze actual markets. The initial frameworks by our theory colleagues made assumptions which ensured that the

1. state variables evolve as a Markov process,
2. and the equilibrium was some form of Markov Perfection (no agent has an incentive to deviate at any value of the state variables).

In these models firms chose ”dynamic controls” ; investments that determine the likely evolution of their state variables. Implicit in the second condition above is that players' have perceptions of the controls' likely impact on the evolution of the state variables (their own and those of their competitors) and through that on their current and future profits, and that these perceptions are consistent with actual behavior (by nature, as well as by their competitors). The standard references here are Maskin and Tirole (1988a and b) for the equilibrium notion and Ericson and Pakes (1995) for the framework brought to applied work. Though, as we will see, there were a number of ways that this framework was successful, it has not had nearly the impact on empirical work that the static framework has, and I want to explore why[2].

2 The Assumptions of the Dynamic Framework.

We start by examining the two assumptions above in the context of symmetric information Markov Perfect models, the first dynamic models to be brought to data.

The Markov Assumption. Except in situations involving active experimentation and learning (where policies are transient), applied work is likely to stick with the assumption that states evolve as a (controlled) time homogeneous Markov process of finite order. There are a number of reasons for this. First the Markov assumption is convenient and fits the data well in the sense that conditioning on a few past states (maybe more than one period in the past) is often all one needs to predict the controls. Second we can bound the gains from unilateral deviations from the Markov assumption (see Ifrach, Weintraub, 2014), and have conditions which insure those deviations can be made arbitrarily small by letting the length of the kept history grow (see White and Scherer, 1994).

Finally, but perhaps most importantly, realism suggests information access and retention conditions as well as computational constraints limit the variables agents actually use in determining their strategies. I come back to this below, as precisely how we limit the memory has implications for the difference between the conditions empirical work can be expected to impose and those most theory models abide by.

[2]There are, of course, some structural dynamic papers that are justifiably well known, see for example Benkard (2004), Collard-Wexler (2013), and Kalouptsidi (2014).

Perfection. The type of rationality built into Markov Perfection is more questionable. It has clearly been put to good use by our theory colleagues, who have used it to explore possible dynamic outcomes in a structured way. It has also been put to good use as a guide to choosing covariates for empirical work which needed to condition on the impacts of dynamic phenomena (e.g. conditioning on the selection induced by exit in the analysis of productivity in Olley and Pakes, 1996). However it has been less successful as an explicit empirical model of agents choices that then combine to form the dynamic response of markets to changes in their environment. This because for many industries it became unwieldly when confronted with the task of incorporating the institutional background needed for an analysis of dynamic outcomes that many of us (including the relevant decision makers) would be willing to trust. The "unwieldlyness" resulted from the dimension of the state space that seemed to be needed (this included at least Maskin and Tirole's, 1988, "payoff relevant" states, or the determinants of the demand and cost functions of each competitor), and the complexity of computing equilibrium policies. The difficulties with the Markov Perfect assumption became evident when we tried to use the Markov Perfect notions to structure

- the estimation of parameters, or to
- compute the fixed points that defined the equilibria or rest points of the system.

The initial computation of equilibrium policies in Pakes and McGuire (1994) discretized the state space and used a "synchronous" iterative procedure. The information in memory allowed the analyst to calculate policies and value functions for each possible state. An iteration circled through each state in turn and updated first their policies to maximize the expected discounted value of future net cash flow given the competitors' policies and the current firm's value function from the last iteration (i.e. it used "greedy" policies given the information in memory), and then updated the values the new policies implied. The test for equilibrium consisted of computing a metric in the difference between the values at successive iterations. If the metric was small enough we were close enough to a fixed point, and the fixed point satisfied all of the equilibrium conditions. The computational burden of this exercise varied directly with the cardinality of the discretized state space which grew (either exponentially or geometrically, depending on the problem) in the number of state variables.

At least if one were to use standard estimation techniques, estimation was even more computationally demanding, as it required a "nested fixed point" algorithm. For example a likelihood based estimation technique would require that the researcher compute equilibrium policies for each value of the parameter vector that the search algorithm tried in the process of finding the maximum to the likelihood. The actual number of equilibria that would need to be calculated before finding the maximum would depend on the problem but could be expected to be in the thousands.

The profession's initial response to the difficulties we encountered in using the Markov Perfect assumptions to structure empirical work was to keep the equilibrium notion and develop techniques to make it easier to circumvent the estimation and computational problems that the equilibrium notion generated. There were a number of useful contribution in this regard. Perhaps the most important of them were:

- The development of estimation techniques that circumvent the problem of repeatedly computing equilibria when estimating dynamic models (that do not require a nested fixed point algorithm). These used non-parametric estimates of the policy functions (Bajari, Benkard, and Levin, 2007), or the transition probabilities (Pakes, Ostrovsky, and Berry, 2007), instead of the fixed point calculation, to obtain the continuation values generated by any particular value of the parameter vector.

- The use of approximations and functional forms for primitives which enabled us to compute equilibria quicker and/or with less memory requirements. There were a number of procedures used; Judd's (1998) book explained how to use deterministic approximation techniques, Pakes and McGuire (2001) showed how to use stochastic algorithms to alleviate the computational burden, and Doraszelski and Judd (2011) showed how the use of continuous time could simplify computation of continuation values.

As will be discussed in several places below many of the ideas underlying these developments are helpful in different contexts. Of particular interest to this paper, the new computational approaches led to an expansion of computational dynamic theory which illuminated several important applied problems. Examples include; the relationship of collusion to consumer welfare (Fershtman and Pakes, 2000), the multiplicity of possible equilibria in models with learning by doing (Besanko, Doraszelski, Kryukov and Satterthwaite, 2010), and dynamic market responses to merger policy (Mermelstein, Nocke, Satterthwaite, and Whinston 2014). On the other hand these examples just sharpened the need for empirical work as the results they generated raised new, and potentially important, possible outcomes from the use of different policies and we needed to determine when these outcomes were relevant. That empirical work remained hampered by the complexity of the analysis that seemed to be required were we to adequately approximate the institutional environment; at least if we continued to use the standard Markov Perfect notions.

2.1 The Behavioral Implications of Markov Perfection.

I want to emphasize the fact that the complexity of Markov Perfection not only limits our ability to do dynamic analysis of market outcomes, it also leads to a question of whether some other notion of equilibria will better approximate agents' behavior. That is the fact that Markov Perfect framework becomes unwieldly when confronted by the complexity of real world institutions, both limits our ability do empirical analysis of market dynamics and raises the question of whether some other notion of equilibria will better approximate agents' behavior. One relevant question then is, if we abandon Markov Perfection can we both

- better approximate agents' behavior and,
- enlarge the set of dynamic questions we are able to analyze?

It is helpful to start by examining why the complexity issue arises. When we try to incorporate what seems to be essential institutional background into our analysis we find that agents are required to both: (i) access a large amount of information (all state variables), and (ii) either compute or learn an unrealistic number of strategies (one for each information set). To see just how demanding this is consider markets where consumer, as well as producer, choices have a dynamic component. This includes pretty much all markets for durable, experience and network goods - that is it includes much of the economy.

In a symmetric information Markov Perfect equilibrium of, say, a durable good market, both consumers and producers would hold in memory at the very least; (i) the Cartesian product of the current distribution of holdings of the good across households crossed with household characteristics, and (ii) each firm's cost functions both for the production of existing products and for the development of new products. Consumers would hold this information in memory, form a perception of the likely product characteristics and prices of future offerings, and compute the solution to a sequential single agent dynamic programming problem to determine their choices. Firms would use the same state variables, take consumers decisions as given, and compute their equilibrium pricing and product development strategies. Since these strategies would not generally be consistent with the consumer's perceptions of those strategies that determined the consumers' decisions, the strategies would then have to be communicated back to consumers who would then have to recompute their value functions and choices based on the updated firm strategies. This process would need to be repeated until we found a "doubly nested" fixed point to the behavior of the agents; until we found strategies where consumers do the best they can given correct perceptions of what producers would do and producers do the best they can given correct perceptions on what each consumer would do. It is hard to believe that this is as good an approximation to actual behavior as the social sciences can come up with.

A Theory "Fix". One alternative to assuming agents know all the information that would be required in a symmetric information Markov Perfect equilibrium is to assume agents only have access to a subset of the state variables. Since agents presumably know their own characteristics and these tend to be persistent, a realistic model would then need to allow for asymmetric information. In that case use of the "perfectness" notion would lead us to a "Bayesian" Markov Perfect solution. Though this will likely reduce information access and retention conditions, it causes a substantial increase in the burden of computing optimal strategies (by either the agents or the analyst). The additional burden results from the need to compute posteriors, as well as optimal policies; and the requirement that they be consistent with one another and hence with equilibrium strategies. The resulting computational burden would make it all but impossible to actually compute optimal policies (likely for many years to come). Of course there is the possibility that agents might learn these policies, or at least policies which maintain some of the logical features of Bayesian Perfect policies, from combining data on past behavior with market outcomes.

Learning Equilibrium Policies. Given its importance in justifying the use of equilibrium policies, there is surprisingly little empirical work on certain aspects of the learning process. There are at least three objects the firm need to accumulate information on; the primitives, the likely behavior of their competitors, and market outcomes given primitives, competitor behavior, and their own policies. There has been empirical work on learning about primitives[3], but very little empirical (in contrast to lab experimental or theoretical) evidence on how firms formulate their perceptions about either their competitors' behavior, or about the impact of their own strategies given primitives and the actions of competitors.

An ongoing study by U. Doraszelski, G. Lewis and myself (2014) delves into these questions. We study the British Electric Utility market for frequency response. Frequency response gives the Independent System Operator (a firm by the name of "National Grid") the ability to keep the frequency of the electricity network within regulated safety bounds. Until November 2005 frequency response was obtained by fiat through a regulation that required all units to allow National Grid to take control of a certain portion of their generating capacity. Starting in November 2005 a monthly auction market for frequency response replaced the regulatory requirement. We have data on bids, acceptances, and auxiliary information on this market from November 2005 until 2012. Note that when this market started the participants had no information available on either competitors' past bids, or about the response of price and quantities to the firms' own bids conditional on the competitors' bids. However they had dealt with the exogenous demand and supply characteristics of this market (monthly variation in demand, prices of fuel,...) for some time.

The results from that study which we are reasonably confident about and have relevance for this paper are that; (i) the bids do eventually converge to what looks like an equilibrium, (ii) after an initial stage where the learning process was too complex for our simple models to approximate adequately, bids for this good converge (and since the good is nearly homogeneous, there is a consequent dramatic fall in the inter-firm variance in bids), and (iii) the many smaller changes in the environment thereafter do not seem to lead to further experimentation. Unfortunately I have little to say about modeling periods of active experimentation as seems to have occurred in the period just after this

[3]See, for e.g. Crawford and Shum (2005), or for a recent contribution and a review of earlier work see Covert (2014).

market was formed. However I will come back to the issue of learning models that do not involve experimentation below.

I now turn to a notion of equilibrium equilibrium that is less demanding than Markov Perfect for both the agents, and the analyst, to use. As we shall see many of the computational and estimation ideas that were developed for Markov Perfect models can be used with the new equilibrium notion, but new issues do arise. In particular, as is explained below, the notion of equilibrium that we propose admits a greater multiplicity than standard Markov Perfect models allow, so we will consider realistic ways of restricting the equilibria analyzed. The last section of the paper uses a computed example of our equilibria to explore computational issues that are associated with it.

3 Less Demanding Notions of Equilibria.

I begin by considering conditions that would be natural candidates for "rest points" to a dynamic system. I then consider a notion of equilibrium that satisfies those, and only those, conditions. The next subsection introduces an algorithm designed to compute policies that satisfy these equilibrium conditions. The algorithm can be interpreted as a learning process. So the computational algorithm could be used to model the response to a change in the industry's institutions, but only changes where it is reasonable to model responses to the change with a simple reinforcement learning process. In particular, I do not consider changes that lead to active experimentation.

Focusing on the equilibrium, or the rest point, makes the job of this subsection much easier. This because strategies at a rest point likely satisfy a Nash condition of some sort; else someone has an incentive to deviate. However it still leaves open the question of the form and purview of the Nash condition. The conditions that I believe are natural and should be integrated into our modeling approach are that

1. agents perceive that they are doing the best they can conditional on the information that they condition their actions on, and that

2. if the information set that they condition on has been visited repeatedly, these perceptions are consistent with what they have observed in the past.

Notice that I am not assuming that agents form their perceptions in any "rational" (or other) way; just that they are consistent with what they have observed in the past, at least at conditioning sets that are observed repeatedly. Nor am I assuming that agents form perceptions of likely outcomes conditional on all information that they have access to. The caveat that the perceptions must only be consistent with past play at conditioning sets that are observed repeatedly allows firms to experiment when a new situation arises. It also implicitly assumes that at least some of the conditioning information sets are visited repeatedly; an assumption consistent with the finite state Markov assumption that was discussed above and which I will come back to below.

I view these as minimal conditions. It might be reasonable to assume more than this, for example that agents know and/or explore properties of outcomes of states not visited repeatedly. Alternatively it might the case that there is data on the industry of interest and the data indicate that behavior can be restricted further. I come back to both these possibilities after a more formal consideration of the implications of the two assumptions just listed.

Formalizing the implications of our two assumptions. Denote the information set of firm i in period t by $J_{i,t}$. $J_{i,t}$ will contain both public (ξ_t) and private ($\omega_{i,t}$) information, so $J_{i,t} = \{\xi_t, \omega_{i,t}\}$. The private information is often information on production costs or investment activity (and/or its outcomes). The public information varies considerably with the structure of the market. It can contain publicly observed exogenous processes (e.g. information on factor price and demand movements), past publicly observed choices made by participants (e.g. past prices), and whatever has been revealed over time on past values of $\omega_{i,t}$.

Firms chose their "controls" as a function of the information at their disposal, or $J_{i,t}$. Typically potential entrants will chose whether or not to enter and incumbents will chose whether to remain active and if they do remain active, how much to invest (in capital, R&D, advertising, ...). Denote the policy chosen by firm i in period t by $m_{i,t} \in \mathcal{M}$, and for simplicity assume that the number of feasible actions, or $\#\mathcal{M}$, is finite (one can deal with continuous values of the control as do Ericson and Pakes (1995); see Fershtman and Pakes (2012)).

Also for simplicity assume we are investigating a game in which firms invest in their own privately observed state ($\omega_{i,t}$) and the outcomes depend only on its own investment choices (not on the choices of its competitors)[4]. In these games the evolution of the firm's own state is determined by a family of distributions which determine the likelihood of the firm's state in the next period conditional on its current state and the amount it invests, or

$$\mathcal{P}_\omega \equiv \{P(\cdot|\omega, m);\ \omega \in \Omega_\omega, m \in \mathcal{M}\}. \tag{1}$$

We assume the number of possible elements in Ω_ω or its cardinality (which will be denoted by $\#\Omega$) is finite (though one can often derive this from primitives, see Ericson and Pakes, 1995).

The firm's choice variables evolve as a function of $J_{i,t}$, and conditional on those choices, the private state evolves as a (controlled) Markov process. This implies that provided the public information evolves as a Markov process, the evolution of $J_{i,t}$ is Markov. In our computational example (section 3), which is about maintenance decisions of electric utility generators, firms observe whether their competitors bid into the auction in each period (so the bids are public information), but the underlying cost "state" of the generator is private information and it evolves stochastically. Here I am simply going to assume that the public information, ξ_t, evolves as a Markov process on Ω_ξ and that $\#\Omega_\xi$ is finite.

In many cases (including our example) the finite state Markov assumption is not obvious. To derive it from primitives we would have to either put restrictions on the nature of the game (see the discussion in Fershtman and Pakes (2012)), or invoke "bounded rationality" type assumptions. I will come back to a more detailed discussion of this assumption below. This because the finite state Markov chain assumption is an assumption I need, and one that can be inconsistent with more demanding notions of equilibrium. For what is coming next one can either assume it was derived from a series of detailed assumptions, or just view it as an adequate approximation to the process generating the data.

Equation (1) and our assumption on the evolution of public information, im-

[4]Though the outcomes could depend on the exogenous processes with just notational changes. This assumption, which generates games which are often referred to as capital accumulation games, is not necessary for either the definition of equilibrium, or the computational and estimation algorithms introduced below. Moreover, though it simplifies presentation considerably, there are many I.O. applications where it would be inappropriate. Consider, for example, a repeated procurement auction for, say timber, where the participants own lumber yards. Their state variable would include the fraction of their processing capacity that their current timber supply can satisfy. The control would be the bid, and the bids of others would then be a determinant of the evolution of their own state. For an analysis of these situations using the notion of equilibrium proposed here see Asker, Fershtman, Jeon, and Pakes (in process).

ply that $J_{i,t}$ evolves as a finite state Markov process, on say $\mathcal{J}$, and that $\#\mathcal{J}$ is finite. Since agents choices and states are determined by their information sets, the "state" of the industry, which we label as s_t, is determined by the collection of information sets of the firms within it

$$s_t = \{J_{1,t}, \dots, J_{n_t,t}\} \in \mathcal{S}.$$

If we assume that there are never more than a finite number of firms ever active (another assumption that can be derived from primitives, (see Ericson and Pakes, 1995), the cardinality of $\mathcal{S}$, or $\#\mathcal{S}$, is also finite. This implies that any set of policies will insure that s_t will wander into a recurrent subset of $\mathcal{S}$, say $\mathcal{R} \subset \mathcal{S}$, in finite time, and after that $s_{t+\tau} \in \mathcal{R}$ with probability one forever (Freedman, 1971). The industry states that are in $\mathcal{R}$, and the transition probabilities among them, will be determined by the appropriate primitives and behavioral assumptions for the industry being studied.

For applied work, it is important to keep in mind from that in this framework agents are not assumed to either know s_t or to be able to calculate policies for each of its possible realizations. Agent's policies (the exit and investment decisions of incumbents, and the entry and investment decisions of potentials entrants) are functions of their $J_{i,t} \in \mathcal{J}$ which is lower dimensional then $s_t \in \mathcal{S}$.

Back to our behavioral assumptions. Our first assumption is that agents chose the policy (the $m \in \mathcal{M}$) that maximizes its own perception of the expected discounted value of future net cash flow. So we need notation for the agent's perceptions of the expected discounted value of future net cash flow that would result from the actions it could chose. The perception of the discounted value from the choice policy m at state J_i will be denoted

$$W(m|J_i), \;\; \forall m \in \mathcal{M} \;\; \& \;\; \forall J_i \in \mathcal{J}.$$

Our second assumption is that at least for J_i that are visited repeatedly, that is for J_i which is a component of an $s \in \mathcal{R}$, the agents' perceptions of these values are consistent with what they observe. So we have to consider what agents observe. When they are at J_i in period t they know the associated public information (our ξ_t) and observe the subsequent public information, or ξ_{t+1}. So provided they visit this state repeatedly they can compute the distribution of ξ_{t+1} given ξ_t. Assuming it is a capital accumulation game and that they know the actual physical relationship between investment and the probability distribution of outcomes (our $\mathcal{P}_\omega$), they can also construct the distribution of of ω_{t+1} conditional on ω_t and m. Together this gives them the distribution of their next period's state, say J_i', conditional on J_i and m. Letting a superscript e denote an empirical distributions (adjusted for the impacts of different m), the conditional distributions are computed in the traditional way, that is by

$$\left\{ p^e(J_i^{'}|J_i, m) \equiv \frac{p^e(J_i^{'}, J_i, m)}{p^e(J_i, m)} \right\}_{J_i^{'}, J_i}.$$

A firm at J_i which choses policy m will also observe the profits it gets as a result of its choice. For simplicity we will assume that the profits are additively separable in m, as would occur for example if profits were a function of all observed firms' prices and m was an additive investment cost. Then once the firm observes the profits it obtains after choosing m it can calculate the profits it would have earned from choosing any $m \in \mathcal{M}$. The empirical distribution of the profits it earns from playing m then allows the firm to form an average profit from playing any m at J_i. We denote those average profits by

$$\pi^e(J_i|m) \;\; \forall m \in \mathcal{M} \;\; \& \;\; \forall J_i \in \mathcal{J}.$$

Note that the profits that are realized at J_i when playing m depend on the policies of (in our example the prices chosen by) its competitors. These in turn depend on its competitors states. In reality there will be a distribution of competitors states, say J_{-i}, when the agent is at J_i, say

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

so in reality the actual expected profits of a firm who plays m at J_i is

$$\pi^e(J_i|m) = \sum_{J_{-i}} \pi(J_i, J_{-i}) p^e(J_{-i}|J_i) - m.$$

Given this notation, our two equilibrium conditions can be formalized as follows.

- If $m^*(J_i)$ is policy chosen at J_i, our first equilibrium condition (i.e. that each agent choses an action which maximizes its perception of its expected discounted value) is written as

$$W(m^*(J_i)|J_i) \geq W(m|J_i), \ \ \forall m \in \mathcal{M} \ \ \& \ \ \forall J_i \in \mathcal{J}. \tag{2}$$

Note that this is an equation on optimal choices of the agent, but provided the agent can learn the $\{W(\cdot|\cdot)\}$ (see below) the agent can make that choice without any information on the choices made by its competitors (the choice becomes analogous to that of an agent playing against nature).

- The second equilibrium condition is that for states that are visited repeatedly (are in $\mathcal{R}$) these perceptions are consistent with observed outcomes. Since $W(m|J_i)$ is the perception of the expected discounted value of future net cash flows, we require that $\forall m$ and $\forall J_i$ which is a component of an $s \in \mathcal{R}$, $W(m|J_i)$ to equal the average profit plus the discounted average continuation value where the distribution of future states needed for the continuation value is the empirical distribution of those states, that is

$$W(m|J_i) = \pi^e(m|J_i) + \beta \sum_{J_i^{'}} W(m^*(J_i^{'})|J_i^{'})p^e(J_i^{'}|J_i). \tag{3}$$

Restricted Experience Based Equilibrium (or REBE). The conditions in equations (2) and (3) above are the conditions of a REBE as defined in Fershtman and Pakes (2012)[5].There also is related earlier work on ”self-confirming” equilibrium (see Fudenberg and Levine, 1983) which is similar in spirit but differs in the conditions it imposes.

A Bayesian Perfect equilibrium satisfies the conditions of a REBE, but so do weaker notions of equilibrium. In particular the REBE does not restrict evaluations of states outside of the recurrent class to be consistent with the outcomes that play at those points would generate. As a result the REBE notion of equilibrium admits greater multiplicity than does Bayesian Perfect notions of equilibrium. We return to the multiplicity issue after explaining how to compute a REBE, as once one has the computational procedure clearly in mind, the multiplicity issue and ways of mitigating it can be explained in a transparent way.

[5]In games where the agent can only use past data to calculate $\{\pi^e(J_i|m)\}$ for $m = m^*(J_i)$ and/or $p^e(J_i'|J_i, m)$ for $m = m^*(J)_i$, Fershtman and Pakes (2012) consider weakening the second condition to only require equation (3) to hold at $m = m^*(J_i)$. They call the equilibrium that results from the weaker notion an EBE (without the restricted adjective).

The Equilibrium Conditions and Applied Work. We already noted that agents are not assumed to compute policies on (or even know) all of s_t; they only need policies conditional on J_i. Now note that there is nothing in our equilibrium conditions that forbids J_i from containing less variables then the decision maker has at its disposal. For example, if agents do not have the capacity to either store too much history or to form differing perceptions of expected discounted values for information sets that detail too much history, one might think it is reasonable to restrict policies to be functions of a subset of the information available to the decision maker. This subset may be defined by a length of the history, or a coarser partition of information from a given history. We come back to the question of how the analyst might determine the information sets that agents' policies condition on below.

The second point to note is related. There is nothing in these conditions that ensures that the policies we calculate on $\mathcal{R}$ satisfy all the equilibrium conditions typically assumed in the game theoretic literature. In particular it may well be the case that even if all its competitors formulated their policies as functions of a particular set of state variables, a particular firm could do better by formulating policies based on a larger set. For example in a model with asymmetric information it is often the case that because all past history may be relevant for predicting the competitors' state, all past history will be helpful in determining current policies. Absent finite dimensional sufficient statistics (which for games are hard to find) this would violate the finite state Markov assumption on the evolution of public information. We still, however, could truncate the history and compute optimal policies for all agents conditional on the truncated history, and this would generate a Markov process with policies that satisfy our conditions (2) and (3).

Fershtman and Pakes (2012) discuss this in more detail and consider alternative ways to ensure REBE policies are the best an agent can do conditional on all agents forming policies as functions of the same underlying state space, and Section 3 uses one of these for comparisons[6]. However I view the less restrictive nature of our conditions as an advantage of our "equilibrium" notion, as it allows agents to have limited memory and/or ability to make computations, and

[6]The example focuses on particular restrictions on the formation of policies, but there are many other ways of restricting policies which would generate Markov chains with similar properties. Indeed the papers I am aware of that compute "approximations" to Markov Perfect equilibria can be reinterpreted in this fashion; see for example Benkard et. al. 2008, and Ifrach and Weintraub, 2014, and the literature cited in those articles.

still imposes an appealing sense of rationality on the decision making process. Moreover in empirical work restrictions on the policy functions may be testable (see section 2.2 below)

3.1 Computational Algorithm.

The computational algorithm is a “reinforcement learning” algorithm[7], similar to the algorithm introduced in Pakes and McGuire (2001). I begin by focusing on computational issues and consider the algorithm’s behavioral interpretation thereafter.

From a computational point of view one of the algorithm’s attractive features is that increases in the number of variables that policies can be a function of (which we will refer to as ”state” variables below) does not (necessarily) increase the computational burden in an exponential (or even geometric) way[8]. Traditionally the burden of computing equilibria scales with both (i) the number of states at which policies must be computed, and with (ii) the number of states we must integrate over in order to obtain continuation values. Depending on the details of the problem, both grow either geometrically or exponentially with the number of state variables, generating what is sometimes referred to as a ”curse of dimensionality”. The algorithm described below is designed to get around both these problems.

The algorithm is iterative and iterations will be indexed by a superscript k. It is also “asynchronous”, that is each iteration only updates a single point in the state space. Thus an iteration has associated with it a location (a point in the state space), and certain objects in memory. The iterative procedure is defined by procedures for updating the location and the memory.

The location, say $L^k = (J_1^k, \dots J_{n(k)}^k) \in \mathcal{S}$ is defined as the information sets of agents that are active at that iteration. The objects in memory, say M^k, include; (i) a set of perceptions of the discounted value of taking action m at location J or $\mathcal{W}^k \equiv \{W^k(m|J_i), \forall m \in \mathcal{M}$ and $\forall J \in \mathcal{J}\}$, (ii) a set consisting of the expected profits when taking action m at location J or $\Pi^k \equiv \{\pi^k(m|J_i), \forall m \in \mathcal{M}$ and $\forall J \in \mathcal{J}\}$, and (iii) a number of times each J has been visited prior to the current iteration, which we denote by h^k. So the

[7]For an introduction to reinforcement learning, see Sutton and Barto, 1998.

[8]The number of state variables in a problem is typically the number of firms that can be simultaneously active times the the number of state variables of each firm.

algorithm must update $(L^k, \mathcal{W}^k, \Pi^k, h^k)$.

Exactly how we structure and update the memory will determine the size of memory constraint and the compute time. Here I suffice with a structure that is easy to explain (the most efficient structure is likely to vary with the properties of the model and the computational facilities available). Also for clarity I work with a model with a specific specification for public and private information. I will assume that the private information, or ω, are payoff relevant states (e.g. costs of production), and the public information that is observed at any state is a function of agents' controls, or $b(m(J))$ (in the electric utility example computed below all agents see whether a generator is bid into the market, but only the owner of the generator sees whether maintenance is done on the generators not bid into the market). In addition the agent is assumed to know the primitive profit function, i.e. $\pi(\cdot, b(m_{-i}))$, which can be used to compute counterfactual profits for any set of competitors' controls (any set of $b(m_{-i})$); i.e. agents can compute $\pi(\cdot, m, b(m_{-i}))$ for $m \neq m_i^*$.

Updating the location. The $\{W^k(m|J_{i,k})\}_m$ in memory represent the agent's perceptions of the expected discounted value of future net cash flow that would result from choosing each $m \in \mathcal{M}$. The agent choses that value of m that maximizes these discounted values (it choses the "greedy" policies). That is for each agent we chose

$$m_{i,k}^* = arg \max_{m \in \mathcal{M}} W^k(m|J_{i,k}).$$

Next we take pseudo random draws on outcomes from the family of conditional probabilities in equation (1), conditional on $m_{i,k}^*$ and $\omega_{i,k} \in J_{i,k}$, that is from $P(\cdot|\omega_{i,k}, m_{i,k}^*)$. The outcomes from those draws determine $\omega_{i,k+1}$ which, together with the current bids of all agents (which is the additional public information), determine the $\{J_{i,k+1}\}$ and hence the new location L^{k+1}.

Updating the memory. Updating the number of visits is done in the obvious way. I now describe the update of perceptions, or of $(\Pi^k, \mathcal{W}^k)$. I do so in a way that accentuates the "learning" interpretation of the algorithm. Since we are using an asynchronous algorithm each iteration only updates the memory associated with the initial location of that iteration.

We assume the agent forms, for each hypothetical choice $m \in \mathcal{M}$, an ex post perception of what its profits and value would have been given the observed choices made by other agents (i.e. $b(m_{-i,k})$). Its profits would have been found by evaluating the profit function at the alternative feasible policies conditional on its private state and its competitors choice of policies, that is $\pi(\omega_{i,k}, m, b(m_{-i,k}))$. Similarly the value would have been those profits plus the continuation values that would have emanated from the alternative choices

$$V^{k+1}(J_{i,k}, m) = \pi(\omega_{i,k}, m, b(m_{-i,k})) + \max_{\tilde{m} \in M} \beta W^k(\tilde{m}|J_{i,k+1}(m)), \tag{4}$$

where $J_i^{k+1}(m)$ is what the $k+1$ information would have been had the agent played m and the *competitors played their actual play*. In the example this would require computing their returns from a counterfactual bid given the bids of the other agents.

The agent knows that $b(m_{-i,k})$ is only one of the possible actions its competitors might take when it is at $J_{i,k}$, as the actual action will depend on its competitors' private information, which the agent does not have access to. So it treats $V^{k+1}(J_{i,k}, m)$ as a random draw from the possible realizations of $W(m|J_{i,k})$, and updates $W^k(m|J_{i,k})$ by averaging this realization with those that had been generated from those prior iterations at which the agents' state was $J_{i,k}$. Formally

$$W^{k+1}(m|J_{i,k}) = \frac{1}{h^k(J_{i,k})+1} V^{k+1}(J_{i,k}, m) + \frac{h^k(J_{i,k})}{h^k(J_{i,k})+1} W^k(m|J_{i,k}),$$

or equivalently

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

An analogous formula is used to update expected profits, i.e. for forming $\{\pi^k(m|J_{i,k})\}$, i.e.

$$\pi^{k+1}(m|J_{i,k}) = \frac{1}{h^k(J_{i,k})+1} \pi(m|J_{i,k}, b(m(J_{-i,k}))) + \frac{h^k(J_{i,k})}{h^k(J_{i,k})+1} \pi^k(m|J_{i,k}).$$

This is a simple form of stochastic integration (see Robbins and Monro, 1951). There are more efficient choices of weights for the averaging, as the earlier iterations contain less relevant information then the later iterations, but I do not pursue that further here[9].

Properties of the Algorithm. Before moving to computational properties note that the algorithm has the interpretive advantage that it can be viewed as a learning process. That is agents (not only the analyst) could use the algorithm, or something very close to it, to learn equilibrium policies. This could be important for empirical work as it makes the algorithm a candidate tool for analyzing how agents might change their policies in reaction to a perturbation in their environment[10].

We now consider computational properties of the algorithm. First note that if we had equilibrium valuations we would tend to stay there; i.e. if * designates equilibrium

$$E[V^*(J_i, m^*)|W^*] = W^*(m^*|J_i),$$

so there is a sense in which the equilibrium is a rest point to the system of stochastic difference equations. I do not know of a proof of convergence of reinforcement learning algorithms for (non zero-sum) games. However we provide a computationally convenient test for convergence below, and my experience is that the randomness in the outcomes of the algorithm together with the averaging over past outcomes that it uses typically is enough to overcome cycling problems that seem to be the most frequent (if not the only) manifestation of non-convergence.

As noted algorithms for computing equilibria to dynamic games have two characteristics which generate computational burdens which increases rapidly as the number of state variables increase (and hence can generate a "curse of dimensionality"). One is the increase of the number of points at which values and policies need to be calculated. In the algorithm just described the only states for which policies and values are updated repeatedly are the points in $\mathcal{R}$. The number of points in $\mathcal{R}$, or $\#\mathcal{R}$ need not increase in any particular way, indeed it

[9] Except to note that the simple weights used here do satisfy Robbins and Monro's, 1951, conditions for convergence (the limit of the sum of the weights is infinite, while the limit of the sum of the squared weights is finite). Though those criteria do not ensure convergence in game theoretic situations, all applications I am aware of chose weights that satisfy them.

[10] Note, however, that were our algorithm to be used as tool for analyzing how agents react to a change in their environment one would have to clarify what information each agent has at its disposal when it updates its perceptions and modify the algorithm accordingly. I.e. in the algorithm described above we use all the information generated by the outcomes to all agents to update the perceptions of each agent, and this may not be possible in an actual application.

need not increase at all, with the dimension of the state space. In the problems I have analyzed $\#\mathcal{R}$ does increase with the dimension of the state space, but at most in a linear (rather than geometric or exponential) way (see the discussion in section 3).

The second source of the "curse of dimensionality" as we increase the number of state variables is the increase in the burden of computing the sum over possible future values needed to compute the continuation values at every point updated at each iteration. In this algorithm the update of continuation values is done as a sum of two numbers regardless of the number of state variables. Of course our estimate of continuation values involves simulation error while explicit integration does not. The simulation error is reduced by repeated visits to the point. The advantage of the simulation procedure is that the number of times a point must be visited to obtain a given level of precision in the continuation values does not depend on the dimension of the state space.

A computational burden of our algorithm that is not present in say, the Pakes and McGuire (1994) algorithm, is that after finding a new location, the reinforcement learning algorithm has to search for the memory associated with that location. In traditional synchronous algorithms one simply cycles through the possible locations in a fixed order. The memory and search constraints typically only become problematic for problems in which the cardinality of $\mathcal{R}$ is quite large, and when they are problematic one can augment our algorithm to use functional form approximations such as those used in the "TD Learning" stochastic approximation literature (see Sutton and Barto, 1998).

Convergence and Testing. Though the algorithm does not necessarily converge, Fershtman and Pakes (2012) provide a test for convergence whose computational burden is both small and independent of the dimension of the state space. To execute the test we first obtain a consistent estimate of $\mathcal{R}$. We then compute a weighted sum of squares of the percentage difference between; (i) the actual expected discounted values from the alternative feasible policies and (ii) our estimates of W, at the points in $\mathcal{R}$. The weights are equal to the fraction of times the points in $\mathcal{R}$ would be visited were those policies followed over a long period of time (it is an $L^2\big(P(\mathcal{R})\big)$ norm of the difference at the different points in $\mathcal{R}$, where $P(\mathcal{R})$ is notation for the invariant measure on $\mathcal{R}$).

First note that any fixed estimate of $\mathcal{W}$, say $\tilde{\mathcal{W}}$, generates policies which define a finite state Markov process for $\{s_t\}$[11]. To obtain a candidate for a recurrent class generated by those policies, say $\mathcal{R}(\tilde{\mathcal{W}})$, start at any s, say s_0, and use the policies that $\tilde{\mathcal{W}}$ imply to simulate a sample path, say $\{s_t\}_{t=1}^{T_1+T_2}$. Let $\mathcal{R}(T_1, T_2, \cdot)$ be the set of states visited at least once between $t = T_1$ and $t = T_2$. This discards the points that are only visited during the first T_1 iterations of the algorithm, and keeps those that are visited between T_1 and T_2. Formally one can show that if $(T_1, T_2) \rightarrow (\infty, \infty)$ in a way which insures $T_2 - T_1 \rightarrow \infty$, $\mathcal{R}(T_1, T_2, \cdot)$ will converge to a recurrent class generated by the policies implied by $\tilde{\mathcal{W}}$. An operational way of checking whether any finite (T_1, T_2) couple were large enough is to continue simulating from T_2 to T_3, where say $T_3 - T_2 \approx T_2 - T_1$. Now check to see if the points visited between T_2 and T_3 are contained in $\mathcal{R}(T_1, T_2, \cdot)$.

Note that the policies we associate with $\tilde{\mathcal{W}}$ are optimal by construction; i.e. $m^*(J_i)$ is chosen to maximize $\{\tilde{W}(m|J_I)\}_{m \in \mathcal{M}}$. This brings us to our last equilibrium condition, the requirement that $\tilde{\mathcal{W}}$ is consistent with the actual outcomes from play for points in $\mathcal{R}$; i.e. we need to check whether

$$\tilde{W}(m|J_i) = \tilde{\pi}(m|J_i) + \beta \sum_{J_i'} \tilde{W}(m^*(J_i')|J_i') p^e(J_i'|J_i), \quad \forall m \in \mathcal{M}, \& J_i \subset s \in \mathcal{R},$$

where $\tilde{\pi}(m|J_i)$ is the algorithm's estimate of expected profits.

In principle we could check this condition by direct summation, but that would be computationally burdensome (indeed it would bring the curse of dimensionality back into play). So we now show how to use simulated sample paths to check it. Start at an $s_0 \in \mathcal{R}$ and use the policies generated by $\tilde{\mathcal{W}}$ to forward simulate. At each J_i visited compute perceived values; i.e. compute $V^{k+1}(\cdot)$ as in equation (4). Since we are simulating a recurrent process on its recurrent class the simulation run will visit each J_i in $\mathcal{R}$ repeatedly. Keep track of the average and the sample variance of the simulated perceived values at each point, say

$$\left(\hat{\mu}\Big(\tilde{W}(m(J_i)|J_i)\Big), \hat{\sigma}^2\Big(\tilde{W}(m(J_i)|J_i)\Big) \right).$$

Let $E(\cdot)$ take expectations over the simulated random draws and, for expositional simplicity, omit the index i. Then note that we can compute $\mathcal{T}_{m,J}$, where

[11] Formally we could gather the implied transition probabilities into the Markov matrix, $Q(s', s|\tilde{\mathcal{W}})$ and describe our first step as finding a candidate for a $\mathcal{R}$ that is generated by $Q(s', s|\tilde{\mathcal{W}})$.

$$\mathcal{T}_{m,J} \equiv E\left(\frac{\hat{\mu}(\tilde{W}_{m,J}) - \tilde{W}_{m,J}}{\tilde{W}_{m,J}}\right)^2$$

$$= E\left(\frac{\hat{\mu}(\tilde{W}_{m,J}) - E[\hat{\mu}(\tilde{W}_{m,J})]}{\tilde{W}_{m,J}}\right)^2 + \left(\frac{E\Big[\hat{\mu}(\tilde{W}_{m,J})\Big] - \tilde{W}_{m,J}}{\tilde{W}_{m,J}}\right)^2.$$

$$= \%Var(\hat{\mu}(\tilde{W}_{m,J})) \ + \ \%Bias^2(\hat{\mu}(\tilde{W}_{m,J})).$$

$\mathcal{T}_{m,J}$ is the percentage mean square error in our estimate of the expected discounted value of taking action m when at state J; i.e. it is the sum of the percentage bias and the percentage variance of the estimate.

Let $\mathcal{T}_J \equiv M^{-1}\sum_{m\in\mathcal{M}} \mathcal{T}_{m,J}$, where $M = \#\mathcal{M}$. $\mathcal{T}_J$ is the average percentage mean square error in the evaluation of the actions that can be taken when at J. $\mathcal{T}_J$ is observed, as is f_J, the fraction of visits to J. As the number of simulation draws grows the law of large numbers implies that we can obtain a consistent estimate of the contribution of the variance in the sample paths to $\mathcal{T}_J$. That is

$$\sum_J f_J \frac{1}{M}\sum_{m\in\mathcal{M}}\left(\frac{\hat{\sigma}^2(\tilde{W}_{m,J})}{\tilde{W}^2_{m,J}}\right) - \sum_J f_J \frac{1}{M}\sum_{m\in\mathcal{M}}\left(\frac{\hat{\mu}(\tilde{W}_{m,J}) - E[\hat{\mu}(\tilde{W}_{m,J})]}{\tilde{W}_{m,J}}\right)^2 \rightarrow_{a.s.} 0.$$

Consequently if

$$Bias(\mathcal{W}_{\mathcal{R}}) \equiv \sum_J f_J \mathcal{T}_J - \sum_l f_J \frac{1}{M}\sum_{m\in\mathcal{M}}\left(\frac{\hat{\sigma}^2(\tilde{W}_{m,J})}{\tilde{W}^2_{m,J}}\right),$$

then

$$Bias(\mathcal{W}_{\mathcal{R}}) \rightarrow_{a.s.} \ \sum_J f_J \frac{1}{M}\sum_{m\in\mathcal{M}}\left(\frac{E\Big[\hat{\mu}(\tilde{W}_{m,J})\Big] - \tilde{W}_{m,J}}{\tilde{W}_{m,J}}\right)^2,$$

an $L^2(\mathcal{P}_{\mathcal{R}})$ norm in the percentage bias, where $\mathcal{P}_{\mathcal{R}}$ is the invariant measure associated with $(\mathcal{R}, \tilde{W})$.

If $Bias(\mathcal{W}_{\mathcal{R}})$ is zero and $\mathcal{R}$ is a recurrent class then all of our equilibrium conditions are satisfied. Notice that this test statistic has an easy interpretation; it is the percentage difference between our estimate of, and the actual expected discounted value of, the net cash flow from the policies that can be undertaken from points in the recurrent class. The test is integrated into the computational algorithm by calling it after every fixed number of iterations, and stopping the algorithm when the estimate of $Bias(\mathcal{W}_{\mathcal{R}})$ is sufficiently small.

3.2 Empirical Challenges and Estimation.

I am going to assume the static profit function is known, as there has been a large literature devoted to empirically analyzing its components (see, for e.g., the first two sections of Ackerberg et. al., 2007, and the literature cited there). The empirical researcher will still need to determine J_i and possibly estimate "dynamic" parameters (parameters that are not determinants of the static profit function).

Determining $\mathbf{J_i}$. In most empirical work the authors simply assume knowledge of J_i, or of the arguments of the policy functions. However given that part of our motivation is to reduce the complexity of the problem by limiting the content of J_i, some discussion of how to determine J_i is in order.

The first thing to note is that what we need to find out is the determinants of the dynamic controls (investment, entry, and exit in our example). In particular it may well be the case that decision makers do not condition on all the information available to them in making these decisions; possibly because making predictions for too fine a partition of the state space is too complicated. As a result specifying a J_i which includes all the information that we know the decision maker has access to may not be necessary or even appropriate.

This suggests two, hopefully reinforcing, methods of determining J_i. The first is an empirical analysis of the determinants of the dynamic controls. The second, which may not always be possible, is to ask decision makers from the industry what their decisions on the dynamic controls depend upon (see for e.g. Wollmann 2015). There are likely to be two sources of error or disturbances in our predictions for the dynamic controls: (i) a "structural" disturbance which results from a determinant of the agent's choice that we do not observe and (ii) a disturbance due to measurement error. Ideally the structural error would be independently distributed over time. and the measurement error component should not be correlated with variables which are thought to be correctly measured. As a result a test of whether the disturbance we obtain from our predictions for the controls satisfies these ideal conditions is that they be uncorrelated (actually independent) of past values of correctly measured variables. If their is indication that the disturbance has a noticeable correlation with past values of correctly measured variables, one should allow for a serially correlated unobserved state variables (see below for further discussion).

Estimating Dynamic Parameters. The estimates needed will be obtained from firm or establishment level data (depending on the parameters being estimated). As a result they will often be based on data sets of similar size as the data sets used in estimating "static" models. These are frequently large enough to obtain reasonably precise parameter estimates (see section 1).

Typically many of the dynamic parameters can be estimated by careful analysis of the relationship between observables without using any of the constructs that need to be computed from the equilibrium to the dynamic model (such as expected discounted values). For example if investment (our control, or m) is observed and directed at improving a measure of a stock of some form (our ω), and the stock is either observable or can be backed out of the profit function analysis, the parameters of $P(\cdot|m, \omega)$ can be estimated directly from the relationship between ω's in adjacent periods and m. However there often are some parameters that can only be estimated through their relationship to perceived discounted values (sunk and fixed costs often have this feature). Also, where possible, more efficient estimators of dynamic parameters that can be estimated without using discounted values can be obtained by using these values.

There is a review of the literature on estimating parameters using the implications of the dynamic model in the third section of Ackerberg et. al. (2007). That review focuses on symmetric information Markov Perfect models and emphasizes the tradeoff between statistical efficiency (in the sense of lower asymptotic variance of an estimated parameter), and computational efficiency (or the computational burden of the estimator), in the choice of estimators. It assumes the state variables of the problem are known to the analyst and provides details on estimators which use them but avoid nested fixed point algorithms[12]. These estimators are all two-step estimators. The first step obtains non-parametric estimates of either: (i) the probabilities of various actions (the "dynamic" controls) as functions of the state variables of the model (typically this includes the probability of entry and exit and a distribution for investment policies; see Bajari Benkard and Levin 2007), or (ii) direct estimates of the Markov transition matrix for the state variables derived from those policies (Pakes Ostrovsky and Berry, 2007). The second step then uses the transitions implied by the non-parametric estimates and the profit function to compute the discounted value of alternative actions conditional on the parameter of interest. It then finds that

[12] Nested fixed point estimators are estimators that require the analyst to compute a new equilibrium every time one evaluates a different parameter vector in the estimation algorithm; see Rust,1994.

value for the parameter vector that makes the prediction for the optimal value of the control as close as possible to the choices actually made for that control.

For example given the profit function, the evolution of the state variables, and the probabilities of exit at each state, we can compute the expected discounted value of an entrant in any period. If the model is correct and we observe entry the expected discounted value generated by entering should have been higher than the sunk cost of entry, whereas if we do not observe entry this expected discounted value should be lower than those costs. Since the average of the realized discounted values should approximate the average of the expected discounted values, the average of the discounted values in the periods when we do, and when we do not, observe entry can be used to estimate bounds on the sunk cost of entry. At the cost of a slight increase in computational burden, one can incorporate heterogeneity in sunk costs and use point (instead of set) estimators in these models (see Ackerberg, Benkard, Berry and Pakes, 2007). Given J_i estimators for Markov Perfect models with asymmetric information can be computed in analogous ways.

In addition I now describe a "perturbation" estimator, similar to the Euler equation estimator for single agent dynamic problems proposed by Hansen and Singleton (1982). This estimator does not require the first step non-parametric estimator, and can be used for estimation in models with asymmetric information (these estimators are not available for symmetric information Markov Perfect models, see below). The perturbation estimator uses the inequality condition for equilibrium policies (i.e. that $W(m^*|J_i) \geq W(m|J_i)$) to generate (set) estimators of parameters. As in the literature on estimating parameters from symmetric information dynamic models, we assume that information on the equilibrium values of controls chosen on the recurrent states are available from past play.

Recall that J_i contains both public and private information. Let J^1 have the same public but different private information than J^2. If a firm is at J^1 it knows it could have played $m^*(J^2)$ and its competitors would respond by playing *on the equilibrium path* from J^2.[13] If J^2 is in the recurrent class we will have data on what competitors would have done were the agent to have chosen $m^*(J^2)$. Provided that choice results in outcomes in $\mathcal{R}$, we can simulate

[13]It is the fact that data would not tell us the response to off the equilibrium path behavior for symmetric information Markov Perfect models that makes the perturbation technique inappropriate for estimating parameters based on those models.

a sample path from J^2 using only observed data on equilibrium play in $\mathcal{R}$. The Markov property insures that the simulated path starting from the deviation to $m^*(J^2)$ will intersect the actual observed sample path at a random stopping time with probability one. From that time forward the two paths would generate the same profits. So the difference in discounted net cash flow from the sequence starting at the actual $m^*(J^1)$ and the sequence starting from the deviation (i.e. from $m^*(J^2)$) is just the difference in discounted returns from the period of the deviation to the time when the paths meet; a difference that we can calculate. Since the initial choice of $m^*(J^1)$ was optimal, the conditional expectation of this difference in discounted profits between the simulated and actual path from the period of the deviation to the random stopping time, should, when evaluated at the true parameter vector, be positive. This yields moment inequalities for estimation as in Pakes, Porter, Ho and Ishii (2015), or the alternatives noted in Pakes (2010).

As noted it may well be important to integrate serially correlated unobservables into these estimation routines. Integrating serially correlated unobservables into these procedures can raise additional issues; particularly if the choice set is discrete. There has been recent work on discrete choice models that allow for serially correlated unobservables (see Arcidiano and Miller, 2011, and Pakes and Porter, 2014), but it has yet to be used in problems that involve estimating parameters that determine market dynamics.

3.3 Multiplicity of Equilibrium Policies.

We noted in section 2.1, that REBE conditions admit more equilibria than Bayesian Perfect conditions. To see why partition the points in $\mathcal{R}$ into "interior" and "boundary" points[14]. Points in $\mathcal{R}$ at which there are feasible (but inoptimal) strategies which can lead outside of $\mathcal{R}$ are 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.

Our conditions only ensure that perceptions of outcomes are consistent with the results from actual play at interior points. Perceptions of outcomes for feasible (but non-optimal) policies at boundary points need not be tied down by actual outcomes. As a result differing perceptions of discounted values at points outside of the recurrent class can support different equilibria. This is a major

[14]This partitioning is introduced in Pakes and McGuire, 2001.

reason for the existence of REBE which are not Bayesian Perfect[15].

One can mitigate the multiplicity problem by adding either empirical information or by strengthening the behavioral assumptions. Without going into details we note that in an empirical application the data will contain information on which equilibria has been played. For example if J_i and m^* are observable we will know policies for states in $\mathcal{R}$. This in turn implies inequalities on the equilibrium $\{W(m|\cdot)\}$ which should rule out some equilibria. Moreover if profits are observed or estimated they can be used, together the transition probabilities, to directly compute estimators of $\{W(m|\cdot)\}$ by simulating sample paths. This may well eliminate other equilibria[16].

These sources of information will be less helpful in two important cases; (i) when attempting to analyze counterfactuals (as we often want to do when examining the impacts of policy or environmental changes)[17], or when (ii) computing equilibria in cases where we are willing to specify primitives but do not have historical data. In these and other cases where we need to augment whatever empirical information is available on the choice of equilibria, it may be reasonable to invoke stronger behavioral assumptions. One possibility is to invoke learning rules, like the one in our algorithm, and simulate equilibria using one or more such rules. This is likely to be more helpful in analyzing counterfactual perturbations to a known environment, as then there is a natural initial condition to start the learning process from (the current state of the industry). A second possibility is to impose additional restrictions on the equilibrium concept per se. I turn to this possibility now.

In many cases prior knowledge or past experimentation will endow agents with realistic perceptions of the value of states outside, but close to, the recurrent class. In these cases we will want to impose conditions that ensure

[15]There are other reasons for differences between REBE and Bayes Nash equilibria. For example, as noted above we do not assume that agents necessarily base their decisions on all the information they either have, or could have, access to. Also typically Bayes Nash Equilibria are defined in terms of consistency of perceived probability distributions with actual actions, whereas we are defining the equilibria in terms of consistency of perceived expected values with actual realized values. There can be different probability distributions that lead to the same expectation. Assuming agents wish to maximize expected discounted value they should be indifferent between two distributions with the same expectations, so I do not see this difference as substantive.

[16]However I know of no formal work which provides details on the extent to which the information in a particular data set limits the set of equilibria that could have generated it.

[17]Assuming historical data is available, there are at two different cases here; one is a counterfactual which changes the underlying state space, and one that does not. If the state space is unchanged *and* one assumes that the counterfactual does not change the equilibrium selection mechanism, it would be possible to use historical data to guide the choice of the counterfactual equilibrium.

that the equilibria we compute are consistent with this knowledge. To accommodate this possibility, Asker, Fershtman, Jeon, and Pakes (2014) propose an additional condition on equilibrium play that insures that agents' perceptions of the outcomes from all feasible actions from points in the recurrent class are consistent with the outcomes that those actions would generate. They label the new condition ”boundary consistency” and provide a computational simple test to determine whether the boundary consistency condition is satisfied for a given set of policies. We now formalize that condition.

Let $B(J_i|\mathcal{W})$ be the set of actions at J_i which is a component of $s \in \mathcal{R}$ which could generate outcomes which are not in the recurrent class (so J_i is a boundary point), and $B(\mathcal{W}) = \cup_{J_i \in \mathcal{R}} B(J_i|\mathcal{W})$ be the set of all possible such actions. Then the "Boundary Consistency" is formulated as follows.

Boundary consistency. Let τ index future periods, and consider a fixed estimate of $\mathcal{W}$, say $\tilde{\mathcal{W}}$. Then $\tilde{\mathcal{W}}$ generates boundary consistent policies if $\forall(m, J) \in B(\tilde{\mathcal{W}})$

$$E\left[\pi(m_i, J_i, J_{-i,0}) + \sum_{\tau=1}^{\infty} \delta^{\tau} \pi\Big(m(J_{i,\tau}), m(J_{-i,\tau})\Big) \Big| m_i = m, J_i = J, \tilde{\mathcal{W}}\right] \leq \tilde{W}(m^*|J),$$

where $E[\cdot|J_i, \tilde{\mathcal{W}}]$ takes expectations over the current states of competitor's and future states of all firms conditional on $J_i = J$ and $m_i = m$ using the policies generated by $\tilde{\mathcal{W}}$. ♠

The boundary consistency condition insures that the policies chosen at the boundary points yield higher discounted values than those of other feasible actions if competitors follow the policies in memory at all states (including the states not in the recurrent class but that communicate with a boundary point if some feasible policy is taken).

The test for boundary consistency uses the fact that we have $\mathcal{W}$ estimates in memory for points outside of $\mathcal{R}$. It uses these $\mathcal{W}$ to determine policies at those points and then simulate sample paths from each (m, J_i) in $B(\tilde{\mathcal{W}})$. The null hypothesis states that the value of all sample paths from feasible but non-optimal policies are less than the value of sample paths from optimal play from the same state. Asker, Fershtman, Jeon, and Pakes (2014), show how to formulate a test of this null[18]. The test should rule out equilibria that are supported by percep-

[18]The test statistic is formed by taking a weighted average of the positive parts of the difference between the

tions of play outside of $\mathcal{R}$ that are unrealistic in the sense that they do not accord with the profits to be earned in those locations. The accuracy of the estimates of the sample paths from boundary points will depend on the components of $\mathcal{W}$ that are not associated with points in $\mathcal{R}$. However they are connected to $\mathcal{R}$ through feasible play, and hence may well have been explored by agents and by the computational algorithm we use to compute equilibria.

Ergodicity. There is another type of multiplicity that may be encountered as there may be multiple recurrent classes for a given equilibrium policy vector. A sufficient condition for the policies to generate an ergodic process (a process with a unique recurrent class) is that there is a single state which can be reached from all states (Freedman, 1971). Ericson and Pakes (1995) use this condition together with assumptions on primitives to prove ergodicity for a certain class of Markov Perfect models. However, in our notation those conditions would be a function of $\mathcal{W}$ on all of $\mathcal{S}$, and our estimates of the W at points not in $\mathcal{R}$ are imprecise (which would make it difficult to determine if those conditions are satisfied). Moreover there are cases of interest where multiple separate recurrent classes are likely (see Besanko, Doraszelski, and Kryukov, 2014). Of course data on (or even qualitative knowledge of) the industry structure should help us pick out which (if there are many) recurrent class is appropriate for the problem at hand.

4 Computational Results from an Example.

It is easiest to explain the computational issues in the context of an example, so most of my focus will be on the example in Fershtman and Pakes (2012) which is concerned with the maintenance decisions of electric utility generators.

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 cur-

estimated value of feasible play and of optimal play of the states in $B(\mathcal{W})$, normalized by the variance of that difference. Since this is a statistic formed from moment inequalities (in contrast to moment equalities) the distribution of this statistic does not have a pivotal form and so needs to be simulated. However the critical values for it are relatively easy to simulate and are compared to the actual value of the test statistic to determine whether to accept the null (for details see Asker, Fershtman, Jeon, and Pakes 2014).

rently electricity is not storable and has extremely inelastic demand might lead to sharp price increases in periods of high demand.

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. Generators have to go down for maintenance periodically. 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. The Fershtman-Pakes paper provides a simple example of a REBE to a game that endogenizes maintenance decisions.

Details of the model. The model has two firms. Firm ”S” has three small generators with low start up costs and high marginal costs (they represent “gas fired” generators), and firm ”L” has two large generators with high start up costs and low marginal costs (they represent “coal fired” generators); see Table 1 below. Each generator can bid supply functions into an independent system operator (the ISO). The ISO sums the bid functions, and then intersects the resultant supply curve with demand (which varies by day of the week) to determine a price. That price is paid to all electricity bid in at any price below it (this is a uniform price auction). The generators have constant marginal cost until the capacity listed in Table 1 after which marginal costs are increasing. Each generator also has a productivity (our ω) which is a private information state variable, and decays stochastically with use. The demand curve is inelastic.

Table 1: Model Details

Parameter	Firm L	Firm S
Number of Generators	2	3
Range of ω	0-4	0-4
MC @ $\omega = (0, 1, 2, 3)^*$	(20,60,80,100)	(50,100,130,170)
Capacity at Const MC	25	15
Costs of Maintenance	5,000	2,000

*MC is constant at this cost until capacity and then goes up linearly. At $\omega = 4$ the generator shuts down.

Each period the firm choses among three actions for each of its generators. It can

- bid the generator into the market, which we denote by $m = 2$,
- withhold the generator from the market and use the period to do maintenance on the generator, our $m = 1$, or
- withhold the generator from the market and do not do any maintenance our $m = 0$.

If the generator is bid in we assume, for simplicity, that it always bids in the same supply curve: so the firm's bid function is $b(m_i) : m_i \to \{0, b_i\}^{n_i}$ where b_i is the fixed bid schedule for the generators of firm i, and firm i has n_i generators. Firms do not see whether their competitors do maintenance but they do see their competitor's bids. So m_i is not in the public information but $b(m_i)$ is.

The cost of producing electricity on each firm's generators is private information; it is a function of the productivity of the generator (our $\omega \in \Omega$) and the quantity of electricity the generator produces ($q_{i,t}$). So the cost function, our $c_i(\omega_{i,t}, q_{i,t})$, is increasing in both its arguments. ω increases stochastically with use, but reverts to a starting value if the firm does maintenance. Formally

- $m_{i,r,t} = 0$ implies $\omega_{i,r,t+1} = \omega_{i,r,t}$,
- $m_{i,r,t} = 1$ implies $\omega_{i,r,t+1} = \overline{\omega}_{i,r}$ where $\overline{\omega}$ is the restart state, and
- $m_{i,r,t} = 2$ implies $\omega_{i,r,t+1} = \omega_{i,r,t} + \eta_{i,r,t}$ where η is a random "productivity shock" with $P(\eta) > 0 \ for \ \ \eta \in \{0, 1\}$.

If d is demand on that day, f is maintenance cost (our "investment"), the price, or p, is determined by the function, $p = p(b(m_i), b(m_{-i}), d)$, and $q = q(b(m_i), b(m_{-i}), d)$ is the allocated quantity vector. Realized profits for firm i is the sum of its profits from its generators (indexed by r) minus the cost of maintenance, or

$$\pi_i(\omega_i, m_i, b(m_{-i}), d) \equiv \sum_r p_t q_{i,r,t} - \sum_r c_i(\omega_{i,r,t}, q_{i,r,t}) - f_i \sum_r \{m_{i,r,t} = 1\}.$$

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

4.1 Conceptual Issues and Their Computational Analogues.

We assumed "a priori" that the state space was finite. As noted above without further restrictions models with asymmetric information will generally have to have policies that depend on all past history in order to insure the equilibrium is perfect (to ensure that if my competitors' condition on a particular public history of play, I can do no better than to condition on that same public history)[19]. We noted a number of different rationals for restricting the history that agents can condition their play on. Perhaps most telling among them is that agents have limited memory and/or ability to make computations, and as a result do not have the capacity to either store too long a history, and/or to form differing perceptions of expected discounted values for information sets that detail too much history. Additionally it might be the case that the finite state space is rich enough to adequately approximate an infinite state space, so there is very little to gain by incorporating more detailed information sets. The example investigates this latter possibility in the context of the model just described (see also the discussion in Ifrach and Weintraub, 2014).

To determine whether or not we have an adequate approximation to a perfect equilibrium we first need to compute a perfect equilibrium which we can compare to. Fershtman and Pakes (2012) prove that one way to ensure that

[19]To see this in the current model note that whether firm 1 bids in during a particular period depends on whether it thinks firm 2 will bid in since if firm 2 does not bid in the price firm 1 receives for its electricity will be higher. Firm 1's perception of whether firm 2 bids in will depend on the last time firm 2 did not bid in (as this is the only time it could have done maintenance). If we go back to the period of when firm 2 did not bid in, its decision at that time depended on whether it thought firm 1 would bid in, which depended on the time before that at which firm 1 bid in, and so on. This recursion on the importance of past information set can go on indefinitely.

there is a perfect equilibrium with a finite state space is to assume that there is full revelation of information every T periods. The policies generated by this equilibrium condition only on the information revealed in the revelation period, the public information that has accumulated since revelation, and the current private information.

They then compute policies that are functions of: (i) same variables but with revelation occurring at different T, and (ii) coarser partitions of the information set than that used to form perfect equilibrium policies. The policies from (i) are perfect equilibrium policies (just different equilibria for each different T). The policies from the calculations in (ii) are not perfect "equilibrium" policies, but do satisfy the conditions in equations (2) and (3) for the restricted information sets. Fershtman and Pakes (2012) then compare statistics of interest that result from simulations using the different policies.

Table 2 presents summary statistics from calculating equilibria assuming different T. It is clear from the table that, at least in this problem, the policies from the $T = 5$ equilibria generate results which are very similar to those from $T = 6$, but the polices from the $T = 3$ equilibria (or even those from $T = 4$) generate results which do not adequately approximate the results from the $T = 6$ equilibria. This suggests that conditioning on a sufficiently long history (in our case T=5) will generate policies that are sufficiently close approximations to the results from policies based on yet longer histories (though I have not made any attempt to formally prove this supposition).

Table 2: Periodic Full Revelation With Different T

	T=3	T=4	T=5	T=6
Summary Statistics.				
Consumer Surplus ($\times 10^{-3}$) 58,000+	550	572	581	580
Profit B ($\times 10^{-3}$)	393	389	384	383
Profit S ($\times 10^{-3}$)	334	324	322	324
Maintenance Cost B ($\times 10^{-3}$)	25.9	21.6	20.2	19.4
Maintenance Cost S ($\times 10^{-3}$)	12.1	11.8	11.8	11.8
Production Cost B ($\times 10^{-3}$)	230.2	235.3	235.1	234.3
Production Cost S ($\times 10^{-3}$)	230.4	226.9	228.1	229.2

F

Tables 3 and 4 present results from two coarser partitions of information sets than the full information set for the $T = 5$ model. In the columns labeled "Finite History T" the results are from a model where there is no revelation of private information and the competitors only keep track of the information publicly revealed in the last T=5 periods. In the columns labeled "Finite History G" the results are from a model where the agents only keep track of the last time each of its competitor's generators was not bid in.

Table 3: Cardinality of $\mathcal{R}$ and Compute Times

	Finite History		Equilibrium
	G	T	(Full Revelation).
Cardinality of Recurrent Class.			
1. Firm B ($\times 10^6$)	5650	38,202	67,258
2. Firm S ($\times 10^6$)	5519	47,304	137,489
Compute Times per 100 Million Iterations (hours; includes test).			
3. Hours	3:04	11:08	17:14
Hours (100 Million) /Size of Recurrent Class (in thousands).			
4 = 3/(1+2)	.26	.130	.083

Table 3 gives the size of the recurrent class and compute times for the three models. Recall that the only points that are visited repeatedly in the algorithm are in the recurrent class. Table 3 provides both the number of such points and the compute time per 100 million iterations of the algorithm. There are two findings to keep in mind; the size of the recurrent class depends on the fineness of the partition of the information sets, and the compute time per million iterations varies directly with the size of the recurrent class. More precisely compute times per a hundred million iterations are increasing and concave in the number of points in the recurrent class. A similar result was found in Pakes and McGuire (2001) who used an analogous algorithm to compute a sequence of symmetric information equilibria with increasing market sizes and hence increases in the cardinality of the underlying state space[20].The relationship between compute times and the size of the recurrent class is largely the result of the time it takes

[20]For example the maximum number of firms active on recurrent points, an endogenous variable which increased with market size, varied from five to ten in those calculations.

to search for the data in memory associated with a new location; a point I come back to below.

Table 4 compares results of interest generated by policies that are a function of coarser partitions of the information then the information set which generates policies which are "perfect" on the recurrent class. It shows that the partitioning implicit in Finite History T is rich enough to give us an accurate picture of the implications of equilibrium play, while that in Finite History G is not. That is we do not seem to need the partition implicit in the full information set, but we do need a partition of that information set that is "rich enough" to provide an adequate approximation to equilibrium play. Of course the conditioning variables that generate such an approximation is likely to vary from problem to problem.

Table 4: Three Asymmetric Information Models

	Finite History		Equilibrium
	G	T	(Full Revelation)
Summary Statistics.			
Consumer Surplus ($\times 10^{-3}$) 58,000+	270	580	581.5
Profit B ($\times 10^{-3}$)	414	384.7	384.5
Profit S ($\times 10^{-3}$)	439	323.5	322.8
Maintenance Cost B ($\times 10^{-3}$)	28.5	20.0	20.2
Maintenance Cost S ($\times 10^{-3}$)	18.0	11.7	11.8
Production Cost B ($\times 10^{-3}$)	226.8	235.5	235.1
Production Cost S ($\times 10^{-3}$)	254.6	228.4	228.1

The results in Tables 3 and 4 are of both analytic and behavioral interest. They suggest three conjectures: (i) the monotonicity of the compute times in the size of the recurrent class, (ii) the coarser the partition of the state space the smaller the recurrent class, and (iii) coarser partitions that are sufficiently rich can provide adequate approximations to optimal policies. Taken together these well may help explain why the decision makers themselves might partition the information available to them in less detailed ways then they could. It also further emphasizes the question of whether we can investigate the issue of the appropriate conditioning set empirically (a task not attempted in Fershtman and Pakes, 2012).

Computational Methods and Burdens. There are many computational issues left to be explored. Two that seemed important determinants of compute time in the work I have been involved in are; (i) the way information is stored, and (ii) the relationship between initial conditions and the computational burden of the algorithm. For storage we have found that storing the public information with a tree structure and the private information with a hash table conditional on public information worked better than using only one or the other of these two possibilities.

Not surprisingly, we have found that if one starts with high enough values for the initial conditions of the algorithm, that is for the components of $\mathcal{W}$ and Π, the algorithm's iterations will explore almost all possible sample paths[21]. As a result the equilibrium which it eventually generates will typically satisfy the "boundary consistency" condition given in section 2.3. This tends to insure we are not supporting the equilibrium by misperceptions of values at boundary points (though ultimately whether actual equilibria are supported by such misperceptions is an empirical question). On the other hand the higher the initial conditions the longer the compute times before the test in section 2.1 is likely to be satisfied. Further in any given application we can now test whether use of smaller starting values result in an equilibrium which is boundary consistent by using the testing procedure discussed in Asker et. al. (2014). That is we can determine whether any set of policies are supported by unrealistic beliefs on outcomes outside of the recurrent class.

This suggest a number of possibilities for reducing the computational burden. For example it may be efficient to use functional form approximations (at least for points outside of the recurrent class) as has been explored in the operations research literature (see Sutton and Barto, 1998). Alternatively, the results above indicate that it might be helpful to start out by computing policies that satisfy the test in section 2.1 with a coarse partition of the information set. A second step would use those policies as starting values for computing policies for the full (or a finer partition of) the information set.

[21]For the results we discuss below we set $\pi_i^{E,k=0}(m_i, J_i) = \pi_i(m_i, m_{-i} = 0, d, \omega_i)$, and $W^{k=0}(\eta_i, m_i|J_i) = \pi_i(m_i, m_{-i} = 0, d, \omega_i + \eta_i(m_i))/(1-\beta)$. They are based on 500 million iterations and generated an $\mathcal{L}^2(\mathcal{P}(\mathcal{R}))$ norm (i.e., a weighted R^2) over .99995. The $\mathcal{L}^2(\mathcal{P}(\mathcal{R})) \geq .99$ at about 200 and flattened out to the minimum between 250 and 350 million (depending on the run).

5 Conclusion.

This paper is meant as a contribution to the development of empirical models for the dynamics of market interactions. It argues for a more realistic framework for that analysis. A framework that does not require agents to; either acquire and retain excessive amounts of information, or compute or learn excessively complicated strategies. What we do require is that agents do not make consistent errors conditional on the information they do compute their policies as a function of. The hope is that this weakening of the traditional restrictions of equilibrium play enables both a better approximation to agents behavior and an analytically more convenient framework for the analyst to use in empirically analyzing that behavior.

References.

- Ackerberg D, Benkard L, Berry S, and Pakes A (2007) Econometric tools for analyzing market outcomes. In Heckman J, Leamer E (ed) The handbook of econometrics, Amsterdam: North-Holland, p 41714276
- Arcidiano P, Miller R (2011) Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. Econometrica 7(6): 1823-1868
- Asker J, Fershtman C, Jeon J, Pakes A (2014) The competitive effects of information sharing" Harvard University in process
- Bajari, P, Benkard L, Levin J (2007) Estimating dynamic models of imperfect competition. Econometrica 75(5): 1331-1370
- Benkard L, (2004) A dynamic analysis of the market for wide-bodied commercial aircraft. Rev Econ Stud, 71: 581-611
- Benkard L, Weintraub G, Van Roy B (2008) Markov perfect industry dynamics with many firms, Econometrica 79(6): 1375-1411
- Berry S, Levinsohn J, Pakes A (1995) Automobile prices in market equilibrium. Econometrica 63(4): 841-890

- Besanko D, Doraszelski U, Kryukov Y, Satterthwaite M (2010) Learning by doing, organizational forgetting and industry dynamics. Econometrica 78(2): 453-508
- Besanko D, Doraszelski U, Kryukov Y (2014) The economics of predation: What drives pricing when there is learning-by-doing? Am Econ Rev 104(3): 868897
- Collard-Wexler A (2013) Demand fluctuations in the ready-mix concrete industry, Econometrica 81(3): 1003-1037
- Covert T (2014) Experiential and social learning in firms: the case of hydraulic fracturing in the Bakken Shale. University of Chicago, Graduate School of Business, working paper
- Crawford G, Shum M (2005) 'Uncertainty and learning in pharmaceutical demand. Econometrica 73(4): 1137-1173
- Crawford G, Yurukoglu A (2013) The welfare effects of bundling in multichannel television markets. Am Econ Rev 102(2): 643-685
- Doraszelski U, Judd K (2011) Avoiding the curse of dimensionality in dynamic stochastic games. Quant Econ 3(1): 53-93
- Doraszelski U, Lewis G, Pakes A (2014) Just starting out: Learning and price competition in a new market. Harvard University working paper
- Einav L, Jenkins M, Levin J (2012) Contract pricing in consumer credit markets. Econometrica 80(4): 1387-1432
- Ericson R, Pakes A (1995) Markov perfect industry dynamics: A framework for empirical work. Rev Econ Stud 62(1): 53-82
- Fershtman C, Pakes A (2000) A dynamic game with collusion and price wars. Rand J Econ 31(2): 207-236
- Fershtman C, Pakes A (2012) Dynamic games with asymmetric information: A framework for applied work. Q J Econ 127(4): 1611-1661
- Freedman D (1971) Markov chains. Holden-Day, San Francisco (Holden-Day series in probability and statistics)

- Fudenberg D, Levine D (1983) Subgame perfect equilibrium of finite and infinite horizon games. J Econ Theory 31(2): 227-256
- Hansen, L, Singleton K (1982) Generalized instrumental variables estimation of nonlinear rational expectations models. Econometrica 50(5): 1269-1286
- Ifrach B, Weintraub G (2014) A framework dynamic oligopoly in concentrated industries. Columbia University working paper
- Judd K (1998) Numerical methods in economics. MIT Press, Cambridge
- Kalouptsidi M (2014) Time to build and fluctuations in bulk shipping. Am Econ Rev 104(2): 564-608
- Maskin E, Tirole J (1988a) A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs. Econometrica 56(3): 549-569
- Maskin E, Tirole J (1988b) A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and edgeworth cycles. Econometrica 56(3): 571-599
- McFadden D (1989) A method of simulated moments for estimation of discrete response models without numerical integration. Econometrica 57(5): 995-1026
- Mermelstein B, Nocke V, Satterthwaite M, Whinston M (2014) Internal versus external growth in industries with scale economies: A computational model of optimal merger policy," NBER Working Papers 20051, National Bureau of Economic Research, Inc.
- Olley S, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. Econometrica 64(6): 1263-1298
- Pakes A, Pollard D (1989) Simulation and asymptotics of optimization estimators. Econometrica 57(5): 1027-1057
- Pakes A, McGuire P (1994) Computing markov perfect nash equilibrium: Numerical implications of a dynamic differentiated product model. Rand J Econ 259(4): 555-589

- Pakes A, McGuire P (2001) Stochastic algorithms, symmetric markov perfect equilibria, and the 'curse' of dimensionality. Econometrica 69(5): 1261-1281
- Pakes A, Ostrovsky M, Berry S (2007) Simple estimators for the parameters of discrete dynamic games (with entry-exit examples). Rand J Econ 38(2): 373-399
- Pakes A (2010) Alternative models for moment inequalities. Econometrica 78(6): 1783-1822
- Pakes A, Porter J (2014) Moment inequalities for multinomial choice with fixed effects. Harvard University working paper
- Pakes A, Porter J, Ho K, Ishii J (2015) Moment inequalities and their application. Econometrica 83(1): 315-334
- Powell J (1994) Estimation of semiparameteric models. In Engle R, McFadden D (ed) Handbook of Ecoometrics, vol 4 Elsevier Science B.V.
- Robbins H, Monro S (1951) A stochastic approximation method. Ann Math Stat 22(3): 400-407
- Rust J (1994) Structural estimation of markov decision processes. In Engle R, McFadden D (ed) Handbook of Econometrics, Volume IV, Elsevier Science, Amsterdam.
- Sutton R, Barto A (1998) Reinforcement learning: An introduction. MIT Press, Cambridge.
- White C, Scherer W (1994) Finite-memory suboptimal design for partially observed markov decision processes. Oper Res 42(3): 439-455
- Wollmann T (2015) Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. Harvard University working paper