STRUCTURAL DYNAMIC MODELS

PAUL SCHRIMPF
NOVEMBER 30, 2011
UNIVERSITY OF BRITISH COLUMBIA
ECONOMICS 628: TOPICS IN ECONOMETRICS

1. INTRODUCTION

These notes cover dynamic structural models. In these models, agents are forward looking and maximize expected payoffs. The models are structural in that they describe agents' preferences. Agents' preferences will be estimated using the principle of revealed preference.

1.1. Notation.

We will use the same notation as Aguirregabiria and Mira (2010). Time is discrete and indexed by t. Agents are indexed by i. The time horizon, T, may be finite or infinite. $s_{it} \in S$ is a vector of state variables that are known at time t. The state variables could include time to allow for non-stationary models (we must include t in the finite horizon case). The state variables can also include time-invariant individual characteristics. $a_{it} \in A$ is an action chosen at time t. Preferences are additively separable over time and discounted at rate β. That is, preferences over possible sequences of s and a are given by

$$\sum_{j=0}^{\infty} \beta^j U(a_{i,t+j}, s_{i,t+j}).$$

Agent's have rational expectations about the evolution of state variables. State variables follow a controlled Markov process, so that the distribution of $s_{i,t+1}$ given all information at time t only depends on $s_{i,t}$ and $a_{i,t}$. Let $F(s_{i,t+1}|a_{it}, s_{it})$ denote the transition distribution. Each period, the agent chooses a_{it} to maximize expected utility.

$$a_{it} \in \arg\max_{a \in A} \mathrm{E}\left[\sum_{j=0}^{\infty} \beta^j U(a_{i,t+j}, s_{i,t+j}) | a_{it} = a, s_{it}\right].$$

The Bellman equation for this problem is

$$V(s_{it}) = \max_{a \in A} U(a, s_{it}) + \beta \mathrm{E}[V(s_{i,t+1})|a, s_{it}]$$

It will also be useful to define the choice specific value function,

$$v(a, s_{it}) = U(a, s_{it}) + \beta \mathrm{E}[V(s_{i,t+1})|a, s_{it}].$$

In the finite horizon case, the value function, V, exists under very weak conditions, see Rust (1994) (essentially we just need the maximization problem at each time to have a solution). Note that the value functions must depend on time when the horizon is finite. We have made this dependence implicit by including t in s_{it}. We will let

$$\alpha(s) \in \arg\max_{a \in A} U(a, s_{it}) + \beta \mathrm{E}[V(s_{i,t+1})|a, s_{it}]$$

denote the policy function.

1.2. **Existence and properties of value and policy functions.** When T is infinite, the existence of the value function requires some assumptions. The easiest case to prove is when U is bounded and continuous. Let $C(S)$ be the space of bounded continuous function from S to $\mathbb{R}$. $C(S)$ is complete under the sup norm. Then $T : C(S) \to C(S)$ defined by

$$T(f)(s) = \max_{a \in A} U(s,a) + \beta \mathrm{E}[f(s')|a,s]$$

is a contraction mapping. Since $C(S)$ is complete, T has a unique fixed point. It is easy to show that unique fixed point solves the original non-recursive problem, see Rust (1994).

$$V(s_{it}) = \max_{a \in A} \mathrm{E}\left[\sum_{j=0}^{\infty} \beta^j U(a_{i,t+j}, s_{i,t+j}) | a_{it} = a, s_{it}\right].$$

This result is nice, but in typical applications U is not bounded. There are similar results for some unbounded U, see Rust (1994) or Stokey and Lucas (1989). Additionally, we will need the policy function,

$$\alpha(s) \in \arg\max_{a \in A} U(a, s_{it}) + \beta \mathrm{E}[V(s_{i,t+1})|a, s_{it}]$$

to be nonrandom and to be approximable in the sense that $V_n \to V$ implies that $\alpha_n \to \alpha$ where

$$\alpha_n \in \arg\max_{a \in A} U(a, s_{it}) + \beta \mathrm{E}[V_n(s_{i,t+1})|a, s_{it}].$$

This and other properties of α follow from appropriate assumptions about U and A. See Stokey and Lucas (1989) for details.

1.3. **Data.** We have panel data on N individuals, each observed for T_i periods. We observe actions, a_{it}, and a sub-vector of the state variable, x_{it}. The unobserved state variable will be ϵ_{it}, so $s_{it} = (x_{it}, \epsilon_{it})$. We also observe some payoff variable,

$$y_{it} = Y(a_{it}, x_{it}, \epsilon_{it}),$$

that contains information about $U(s,a)$, but is not one of the state variables. For example, y_{it} could be revenues of a firm, or an individual's earnings.

1.4. **Examples.** The following examples are taken from Aguirregabiria and Mira (2010).

Example 1.1 (Retirement)**.** Consider the choice of when to retire. Let $a_{it} = 1$ if an agent is working and $a_{it} = 0$ if retired. Suppose T is the age at death. The payoff function could be

$$U(a_{it}, x_{it}, \epsilon_{it}) = \mathrm{E}[c_{it}^{\theta_1} | a_{it}, x_{it}] \exp\left(\theta_2 + \theta_3 h_{it} + \theta_4 \frac{t}{1+t}\right) - \theta_5 a_{it} + \epsilon(a_{it})$$

where c_{it} is consumption, θ_1 is the coefficient of relative risk aversion, h_{it} is health, and the expression in the exp captures the idea that the marginal utility of consumption could vary with health and age. $-\theta_5 a_{it}$ captures the disutility of working.

Example 1.2 (Entry/exit)**.** A firm is deciding whether to operate in a market. Its per-period profits are

$$U(a_{it}) = a_{it}\left(\theta_R \log(S_t) - \theta_N \log\left(1 + \sum_{j \neq i} a_{jt}\right) - \theta_F - \theta_E(1 - a_{i,t-1}) + \epsilon_{it}\right)$$

where a_{it} is whether the firm operates at time t. S_t is the size of the market, $\sum_{j \neq i} a_{jt}$ is the number of other firms operating. θ_F is a fixed operating cost, and θ_E is an entry cost.

2. IDENTIFICATION

2.1. **General non-identification.** Rust (1994) shows that without some restrictions, the above model is not identified. In the data, we observe

$$\alpha(s) = \arg\max_{a \in A} v(a, s).$$

Rust (1994) calls this the reduced form of a Markov decision problem. The structure of the a Markov decision problem is the mapping, $\Lambda : \{\beta, U, F\} \to \alpha$ defined by

$$\alpha(s) = \arg\max_{a \in A} v(a, s)$$

where

$$v(a, s) = U(a, s) + \beta \mathrm{E}[\max_{a' \in A} v(a', s') | a, s].$$

Our goal is to identify the structure, (β, U, F).

Definition 2.1. Primitives (β, U, F) and $(\tilde{\beta}, \tilde{U}, \tilde{F})$ are **observationally equivalent** if

$$\Lambda(\beta, U, F) = \Lambda(\tilde{\beta}, \tilde{U}, \tilde{F}).$$

It's clear that $\Lambda(\beta, U, F) = \Lambda(\beta, aU + b, F)$, so we can at most identify U up to a linear transformation. Rust (1994) shows that we can identify even less.

Lemma 2.1. *Let f be any measurable function of s. Define:*

$$\tilde{U}_f(a, s) = U(a, s) + f(s) - \beta \mathrm{E}[f(s') | a, s]$$

. Then

$$\Lambda(\beta, U, F) = \Lambda(\beta, \tilde{U}_f, F).$$

Proof. Let $v(a, s)$ be the choice specific value function for (β, U, F). We can guess and verify that

$$\tilde{v}_f(a, s) = v(a, s) + f(s)$$

is the choice specific value function for $(\beta, \tilde{U}_f, F)$.

$$\begin{aligned}
\tilde{v}_f(a, s) =& \tilde{U}_f(a, s) + \beta \mathrm{E}[\max_{a' \in A} \tilde{v}_f(a', s') | a, s] \\
=& U(a, s) + f(s) - \beta \mathrm{E}[f(s') | a, s] + \beta \mathrm{E}[\max_{a' \in A} v(a', s') + f(s') | a, s] \\
=& U(a, s) + f(s) + \beta \mathrm{E}[\max_{a' \in A} v(a', s') | a, s] \\
=& v(a, s) + f(s).
\end{aligned}$$

Since $f(s)$ does not depend on a,

$$\arg\max_{a \in A} v(a, s) = \arg\max_{a \in A} v(a, s) + f(s),$$

so

$$\alpha(s) = \tilde{\alpha}_f(s)$$

and the models are observationally equivalent. □

This lemma shows that given any U, there are many observationally equivalent $\tilde{U}_f$. Even knowing β and F, U is not identified. Rust (1994) also shows that given any reduced form policy function, there is a U that generates it.

Lemma 2.2. *Let $\alpha : S \to A$ and A be discrete. Then $U(a,s) = 1\{a = \alpha(s)\} - \beta$ along with any F and β results in*

$$\Lambda(\beta, U, F) = \alpha.$$

Proof. Guess and verify that $v(a,s) = 1\{a = \alpha(s)\}$. □

This lemma shows that having a Markov decision problem places no restrictions on the observed policy function (other than it being Markovian).

2.2. **Identification in dynamic discrete decision models .** Magnac and Thesmar (2002) expand on the non-identification result of Rust (1994), and give conditions under which identification is possible. Magnac and Thesmar (2002) specifically focus on dynamic discrete decision models that satisfy the following assumptions. These assumptions have been used in most applications.

A1 (Discrete actions)**.** *A is discrete and finite.*

A2 (Additive separability)**.** *Instantaneous utility functions are given by*

$$U(a,x,\epsilon) = u(a,x) + \epsilon(a)$$

where

$$\mathrm{E}[\epsilon(a)|x] = 0$$

for each $a \in A$. The cdf of ϵ, denoted G, is absolutely continuous with respect to Lebesgue measure in $\mathbb{R}^{|A|}$.

A3 (Conditional independence)**.** *For any $t \neq t'$, ϵ_{it} and $\epsilon_{it'}$ are independent conditional on x and a.*

As a result of these assumptions,

$$\begin{aligned} v(a,x,\epsilon) =& U(a,x,\epsilon) + \beta \mathrm{E}[\max_{a \in A} v(a',x',\epsilon')|x,a,\epsilon] \\ =& u(a,x) + \epsilon(a) + \beta \mathrm{E}[\max_{a \in A} \tilde{v}(a,x) + \epsilon(a)|x,a] \\ =& \tilde{v}(a,x) + \epsilon(a) \end{aligned}$$

Magnac and Thesmar (2002) assume that the support of x is discrete and finite as well. However, this assumption is unnecessary, as shown by Bajari, Chernozhukov, Hong, and Nekipelov (2009). In any case, let

$$\mathrm{P}(a|x) = \mathrm{P}\left(a \in \arg\max \tilde{v}(a,x) + \epsilon(a)\right).$$

Hotz and Miller (1993) show that these equations can be inverted to yield

$$\tilde{v}(a,x) - \tilde{v}(0,x) = q(a, \mathrm{P}(\cdot|x); G)$$

where $0 \in A$ is some reference action, and q depends on the distribution of ϵ, G. In many applications, it is assumed that $\epsilon(a)$ has an extreme value distribution,

$$G_a(\epsilon) = \frac{e^{\epsilon}}{1 + e^{\epsilon}}.$$

In this case,

$$\tilde{v}(a,x) - \tilde{v}(0,x) = \log(\mathrm{P}(a|x)) - \log(\mathrm{P}(0|x)).$$

Regardless, we can write

$$\begin{aligned} \tilde{v}(0,x) =& u(0,x) + \beta \mathrm{E}[\max_{a' \in A} \tilde{v}(a',x') + \epsilon(a')|a,x] \\ =& u(0,x) + \beta \mathrm{E}[\max_{a' \in A} \tilde{v}(a',x') - \tilde{v}(0,x') + \epsilon(a')|0,x] + \beta \mathrm{E}[\tilde{v}(0,x')|0,x] \end{aligned}$$

The middle term is some function of G and the observed choice probabilities, say

$$M(x, \mathrm{P}(\cdot|x), G) = \beta \mathrm{E}[\max_{a' \in A} \tilde{v}(a', x') - \tilde{v}(0, x') + \epsilon(a')|0, x].$$

Suppose we normalize $u(0, x) = 0$. Then we have

$$\tilde{v}(0, x) = M(x, \mathrm{P}(\cdot|x), G) + \beta \mathrm{E}[\tilde{v}(0, x')|0, x]$$

If the model is stationary (in particular the supports of x and x' coincide), then it is easy to show that this equation has a unique solution for $\tilde{v}(0, x)$. If the model is not stationary, then in the finite horizon case, we can begin with $t = T$ and set $\tilde{v}(0, x) = M(x, \mathrm{P}(\cdot|x), G)$. In either case, $\tilde{v}(0, x)$ is identified if M, which depends on G and observable data, and β are known.

Lemma 2.3. *Suppose G and β are known, and $u(0, x) = 0$, then $\tilde{v}(0, x)$ is identified.*

Finally, observe that

$$\tilde{v}(a, x) = u(a, x) + \beta \mathrm{E}[\max_{a' \in A} \tilde{v}(a', x') + \epsilon(a')|a, x]$$
$$u(a, x) = \tilde{v}(a, x) - \beta \mathrm{E}[\max_{a' \in A} \tilde{v}(a', x') + \epsilon(a')|a, x]$$

so u is also identified.

Theorem 2.1 (Identification)**.** *If assumptions A1-A3 hold, and G, β, and $u(0, x) = 0$ are known, and the model is infinite horizon and stationary, or finite horizon and we observe all time period, then $u(a, x)$ is identified.*

This theorem shows that assuming errors are additively separable with a known distribution, knowing the discount factor, and normalizing $u(0, x)$ is sufficient for identification. Magnac and Thesmar (2002) show that given assumptions A1-A3, knowing G, β, and $u(0, x)$ is also necessary in that $G' \neq G$ or $\beta' \neq \beta$ or $u(0, x)' \neq u(0, x)$, there exists observationally equivalent $u(a, x)$. It is easy to see this result from our discussion above. Every step that we took was constructive. We explicitly found $u(a, x)$ given β, G, and $u(0, x)$. If we change any of those three things, we will still end up with some $u(a, x)$, but it will be different.

2.3. **Identification in dynamic decision models with continuous actions.** Schrimpf (2011) gives an analogous identification result for dynamic games with continuous actions.

3. Estimation

Given the above identification results, we will begin by focusing on estimating models that satisfy assumptions A1, A2, and A3. We will also assume that the distribution of ϵ, G, is known. Let's also suppose that the payoff function has been parametrically specified, $u(a, x; \theta_u)$, and the transition distribution is also parametrically specified, so that in particular, the conditional expectation of any function of x' and a' given x and a can be written as $\mathrm{E}[\cdot|x, a; \theta_p]$. Given data on a_{it} and x_{it} we want to estimate θ_u and θ_p. To begin with, we will treat θ_u and θ_p as finite dimensional parameters, but we will discuss nonparametric estimation later (i.e. allow θ_p and/or θ_u to be infinite dimensional). Throughout, we will focus on infinite horizon problems.

3.1. **Nested fixed point.** This subsection is largely based on Rust (1994). The nested fixed point algorithm is a maximum likelihood estimator for θ. It is computationally intensive because it requires numerically maximizing the likelihood, and each time the likelihood is evaluated, the value function is solved for. Thus there are nested fixed points being solve. The inner fixed point is the value function, the outer one is the maximization of the likelihood.

3.1.1. *Solving for the value function.* For any given value of $\theta = (\theta_u, \theta_p)$, we can compute the value function by solving the Bellman equation. There are many ways to do this. The most straightforward method is value function iteration. That is, begin with some guess for the value function, say $V_0(x, \epsilon; \theta)$. Then update the guess by setting

$$V_1(x, \epsilon; \theta) = \max_{a \in A} u(a, x; \theta_u) + \epsilon(a) + \beta \mathrm{E}[V_0(x', \epsilon'; \theta) | x, a; \theta_p] \tag{1}$$

Repeat until convergence. This method of computing V is called value function iteration. Value function iteration is stable and globally convergent, but only converges at a geometric rate equal to β.

$$\|V_{i+1} - V\| \lesssim \beta \|V_i - V\|$$

As a result, when β is near one we may need many iterations before convergence.

If x or ϵ are continuous, we cannot represent $V(x, \epsilon; \theta)$ in a computer. Instead, we can only work with some approximation of V. The most common approach is to discretize x and ϵ. That is, divide the support of x and ϵ into a finite number of bins, and approximate V as being constant in each of those bins. Then V can be represented by a vector consisting of the values of $V(x, \epsilon; \theta)$ for each x, ϵ bin. V could also be approximated by a series or kernel.

Once you have chosen an way to approximate V, the Bellman equation 1 becomes a system of non-linear equations. Rather than using value function iteration, you could solve this system of equation using Newton's method. Newton's method locally converges quadratically, so it theoretically requires fewer iterations. That is, if V_i is close enough to V, then

$$\|V_{i+1} - V\| \lesssim \|V_i - V\|^2 .$$

However, Newton's method need not be globally convergent, so if the initial V_0 is far from, V, it can take longer than value function iteration.

One method that is equivalent to Newton's method is policy iteration. In policy iteration, you begin with an initial guess for the policy function, $\alpha_0(x, \epsilon)$. Then you compute the corresponding value function,

$$V_\alpha(x, \epsilon; \theta) = u(\alpha(x, \epsilon), x; \theta_u) + \epsilon(\alpha(x, \epsilon)) + \beta \mathrm{E}[V_0(x', \epsilon'; \theta) | x, \alpha(x); \theta_p].$$

This is much easier than solving the Bellman equation, because there is no maximization. Given a way of approximating V, this equation can often be written as

$$\mathbf{V}_\alpha = \mathbf{U}(\theta_u) + \beta P(\theta_p) \mathbf{V}_\alpha$$

where $\mathbf{V}$ and $\mathbf{U}(\theta_u)$ are vectors and $P(\theta_p)$ is a square matrix. You can then solve for V_α as

$$\mathbf{V}_\alpha = (I - \beta P(\theta_p))^{-1} \mathbf{U}(\theta_u).$$

After solving for V_α, you update the policy function by setting

$$\alpha_1(x, \epsilon) = \arg\max_{a \in A} u(x, a; \theta) + \epsilon(a) + \beta \mathrm{E}[V_{\alpha_0}(x', \epsilon') | x, a],$$

and repeat until convergence. Like Newton's method, policy function iteration locally converges quadratically, but is not globally convergent or stable. In practice, it is often effective to begin with value iteration and then switch to policy iteration.

There are many details that we have left unspecified in this discussion. It is difficult to get a good understanding of how solving dynamic programs works without going through some examples in detail. On homework 6, you will solve a dynamic program by discretizing the state space. On my website, `http://faculty.arts.ubc.ca/pschrimpf/14.170/programming.html`, there is an example that uses series functions to approximate the value and policy functions.

3.1.2. *The likelihood.* Once we have computed $V(x,\epsilon;\theta)$, we can easily compute the likelihood. The probability of a conditional on x is

$$\mathrm{P}(a|x;\theta) = \mathrm{E}_\epsilon \left[1\{a = \arg\max_{\tilde{a}\in A} u(a,x;\theta_u) + \epsilon(a) + \beta \mathrm{E}[V(x',\epsilon';\theta)|\tilde{a},x;\theta_p]\} \right]. \tag{2}$$

Then the log likelihood is

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T_i} \log \mathrm{P}(a_{it}|x_{it};\theta)$$

We estimate θ by maximizing the likelihood. Rust (1994) suggests using either the BHHH [Brendt, Hall, Hall, and Hausman (1974)] or BFGS algorithm to maximize the likelihood. Both of these algorithm are quasi-Newton methods that do not require explicitly calculating the Hessian, but still converge at a faster than linear (although not quite quadratic rate). The BHHH approximates the Hessian using the outer-product of the gradient. From the information equality, we know that this approximation will be exact in an infinite sample. The BFGS algorithm approximates the Hessian by the change in the gradient at various evaluations of the likelihood.

3.1.3. *Two and three step nested fixed point.* The nested fixed point algorithm is computationally intensive. One way to slightly reduce the amount of computation is to first estimate θ_p. Recall that θ_p enters $\mathrm{E}[\cdot|x,a;\theta_p]$. Since x and a are observed, this conditional expectation functional can be estimated without using the full model. Thus, we can first estimate θ_p, and then maximize the likelihood with respect to θ_u only. This maximization should require fewer iterations because it can search over a lower dimensional space. The resulting two-step estimates will not be as efficient as the one-step estimates. However, they will be consistent. As always with consistent estimates, we can then perform one (or more) Newton step(s) of the full likelihood to obtain efficient estimates.

3.2. **Hotz and Miller's CCP method.** Even the two-step version of the nest fixed point algorithm can be computationally infeasible for large problems. Hotz and Miller (1993) propose an estimator that is much easier to compute. Suppose that the payoff function is linear in parameters,

$$u(a,x;\theta_u) = z(a,x)'\theta_u,$$

where $z(a,x)$ is known. Then the choice specific value functions are given by

$$\begin{aligned}
\tilde{v}(a,x) =& z(a,x)'\theta_u + \beta \mathrm{E}\left[\max_{a'\in A} \tilde{v}(a',x') + \epsilon(a)|a,x\right] \\
=& z(a,x)'\theta_u + \beta \mathrm{E}\left[\sum_{a'\in A} v(a',x')\mathrm{P}(a'|x')|a,x\right] + \beta \mathrm{E}[\mathrm{E}[\epsilon(a')|\alpha(x',\epsilon)=a']|a,x] \\
=& z(a,x)'\theta_u + \beta \mathrm{E}\left[\sum_{a'\in A} v(a',x')\mathrm{P}(a'|x')|a,x\right] + \beta \mathrm{E}[\sum_{a'\in A} \mathrm{E}[\epsilon(a')|\tilde{v}(a',x')+\epsilon(a') \geq \tilde{v}(a'',x')+\epsilon(a'')\forall a''\in A]\mathrm{I}
\end{aligned}$$

Let

$$e(a,x) = \beta \mathrm{E}[\sum_{a'\in A} \mathrm{E}[\epsilon(a')|\tilde{v}(a',x')+\epsilon(a') \geq \tilde{v}(a'',x')+\epsilon(a'')\forall a''\in A]|a,x].$$

Then it is clear that we can write

$$\tilde{v}(a,x) = \tilde{z}(a,x)'\theta_u + \tilde{e}(a,x)$$

where

$$\tilde{z}(a,x) = z(a,x) + \beta \mathrm{E}[\sum_{a'\in A} \tilde{z}(a',x')\mathrm{P}(a'|x')|a,x]$$

and

$$\tilde{e}(a,x) = e(a,x) + \beta \mathrm{E}[\sum_{a' \in A} \tilde{e}(a',x')\mathrm{P}(a'|x')|a,x].$$

Note that $\mathrm{E}[\epsilon(a')|\tilde{v}(a',x') + \epsilon(a') \geq \tilde{v}(a'',x') + \epsilon(a'')\forall a'' \in A]|a,x]$ only depends on the distribution of ϵ, G, and $\tilde{v}(a',x') - \tilde{v}(a'',x')$. As in the identification section, Hotz and Miller (1993) show that

$$\tilde{v}(a',x') - \tilde{v}(a'',x') = q(a',a'',\mathrm{P}(\cdot|x),G)$$

for some known function q. $\mathrm{P}(a|x)$ can be estimated from the observed a and x. $\mathrm{E}[\cdot|a,x]$ can also be estimated from the observed a and x. Therefore, estimates of $\tilde{z}$ and $\tilde{e}$ can be formed before estimating θ. Denote these by $\hat{z}$ and $\hat{e}$. Once we have $\hat{z}$ and $\hat{e}$, we can form estimates of the probability of any action conditional on x given any θ.

$$\hat{\mathrm{P}}(a|x;\theta) = \mathrm{P}\left(a = \arg\max_{a' \in A} \hat{z}(a',x)'\theta + \hat{e}(a',x)\right)$$

Hotz and Miller (1993) propose estimating θ by GMM using moment conditions of the form

$$\sum_{i=1}^{N}\sum_{t=1}^{T_i} f(x_{it})\left(1\{a_{it} = a\} - \hat{\mathrm{P}}(a|x_{it};\theta)\right).$$

Aguirregabiria and Mira (2002) show that if you use the pseudo maximum likelihood to estimate θ with pseudo-likelihood function,

$$\tilde{\mathcal{L}}(\theta) = \sum_{i=1}^{N}\sum_{t=1}^{T_i} \log \hat{P}(a_{it}|x_{it};\theta),$$

then $\hat{\theta}$ has the same asymptotic distribution as when you use the two-step nested fixed point estimator. However, Aguirregabiria and Mira (2010) describe Monte Carlo evidence that in finite samples, this pseudo maximum likelihood estimator can have large bias.

- Nested pseudo likelihood.
- Using simulation.

4. DYNAMIC GAMES

The above methods can be applied to dynamic games as well as dynamic decision problems. As above, let's restrict our attention to games with discrete states and actions. Suppose there are N players indexed by i. Each player chooses a discrete action $a_{it} \in A$ given the current observed state $x_t = (x_{1t}, ..., x_{Nt})$ and a private shock ϵ_{it}. ϵ_{it} is only known to player i. x_t is common knowledge among all players. The payoff of player i depends on the actions of all players, $a_t = (a_{1t}, ..., a_{Nt})$, the state, x_t, and the private shock, ϵ_{it}. Assume that payoffs are additively separable in ϵ,

$$U_i(a_t, x_t, \epsilon_{it}) = u_i(a_t, x_t) + \epsilon_{it}(a_{it}).$$

We will assume that ϵ is iid across time and players with CDF G. We also assume that the evolution of x_t does not depend on ϵ, $F(x_{t+1}|a_t, x_t, \epsilon_t) = F(x_{t+1}|a_t, x_t)$.

We restrict our attention to Markov perfect equilibria, so strategies only depend on the current state. Let $\alpha : (X \times \mathbb{R})^N \to A^N$ denote a vector of strategies. α_i is the strategy of player i, and α_{-i} is

the strategy of the other players. Let $V_i^\alpha(x_t, \epsilon_{it})$ be the value function for player i. The associated integrated value function is

$$\begin{aligned}\bar{V}^\alpha(x) &= \int V_i^\alpha(x_t, \epsilon_{it}) dG(\epsilon_{it}) \\ &= \int \left(\max_{a_{it} \in A} v_i^\alpha(x_t, a_{it}) + \epsilon_{it}(a_{it}) \right) dG(\epsilon_{it})\end{aligned}$$

where v_i^α is the choice specific integrated value function, which solves

$$v_i^\alpha(a_{it}, x_t) = \mathrm{E}_{\epsilon_{-i}} \left[u_i(a_{it}, \alpha_{-i}(x_t, \epsilon_{-it}), x_t) + \beta \mathrm{E}_x [\bar{V}_i^\alpha(x_{t+1}) | a_{it}, \alpha_{-i}(x_t, \epsilon_{-it}), x_t] \right] \tag{3}$$

where the outer expectation is over ϵ_{-it} and the inner one is over x_{t+1}. As in the single agent case, we can define conditional choice probabilities,

$$\begin{aligned}\mathrm{P}_i^\alpha(a_i|x) &= \mathrm{P}\left(a_i = \arg\max_{j \in A} v_i^\alpha(j, x) + \epsilon_{it}(j) | x \right) \\ &= \int 1 \left\{ a_i = \arg\max_{j \in A} v_i^\alpha(j, x) + \epsilon_{it}(j) \right\} dG(\epsilon_{it}).\end{aligned}$$

In the single agent case, the restrictions of the model end here. The conditional choice probabilities must be consistent with maximizing the value function. In a dynamic game, we have an additional restriction: the conditional choice probabilities should form an equilibrium. To add this constraint, rewrite the expectation over ϵ_{-i} in (3) as

$$v_i^\alpha(a_{it}, x_t) \equiv v_i^P(a_{it}, x_t) = \sum_{a_{-i} \in A^{N-1}} \mathrm{P}_{-i}(a_{-i}|x_t) \left(u_i(a_{it}, a_{-i}, x_t) + \beta \mathrm{E}_x [\bar{V}_i^\alpha(x_{t+1}) | a_{it}, a_{-i}, x_t] \right) \tag{4}$$

where

$$\mathrm{P}_{-i}(a_{-i}|x) = \prod_{j \neq i}^{N} \mathrm{P}(a_j|x).$$

Let

$$\Lambda(a | v_i^P(\cdot, x_t)) = \int 1 \left\{ a_i = \arg\max_{j \in A} v_i^P(j, x) + \epsilon_{it}(j) \right\} dG(\epsilon_{it}).$$

Then the equilibrium condition is that

$$\mathrm{P}_i(a|x) = \Lambda(a | v_i^P(\cdot, x))$$

or in vector form $\mathrm{P} = \mathbf{\Lambda}(v^P)$ where P is the vector of conditional choice probability functions for all players, and $\mathbf{\Lambda}$ is similarly defined.

Viewed as a function of P, $\mathbf{\Lambda}$ is a mapping from $[0,1]^{N|X|}$ to $[0,1]^{N|X|}$. Moreover, $\mathbf{\Lambda}$ is continuous, and the unit cube in $\mathbb{R}^{N|X|}$ is convex compact set, so by Brouwer's fixed point theorem, there exists at least one equilibrium. There are often multiple equilibria.

4.1. **Identification.** The identification argument for single-agent dynamic decision problems shows that given G, β, and $\mathrm{E}_\epsilon[u(0, \alpha_{-i}(x, \epsilon_{-i}), x_t)] = 0$, we can identify the expectation over other player's actions of the payoff function,

$$\mathrm{E}_\epsilon[u(a_i, \alpha_{-i}(x, \epsilon_{-i}), x)] = \sum_{a_{-i}} \mathrm{P}(a_{-i}|x) u(a_i, a_{-i}, x)$$

When x and a are discrete, the left hand side of this equation takes $|A||X|$ identified values. On the right side, we know $\mathrm{P}(a_{-i}|x)$ is known, but $u(a_i, a_{-i}, x)$ is not, and it can take $|A|^N|X|$ values. Therefore, we need some restriction if want to identify the payoff function. The usual approach

is to assume that there are some state variables that enter the payoff of the other players, but not player i's payoff directly. Then $\mathrm{P}(a_{-i}|x)$ depends on all state variable, but $u(a_i, a_{-i}, x_i)$ does not. Then u is identified if the system of equations

$$\mathrm{E}_\epsilon[u(a_i, \alpha_{-i}(x, \epsilon_{-i}), x)] = \sum_{a_{-i}} \mathrm{P}(a_{-i}|x) u(a_i, a_{-i}, x_i)$$

has a unique solution for $u(a_i, a_{-i}, x_i)$.

4.2. **Estimation.** Each of the estimation approaches described for single agent problems above can also be used for dynamic games. As before, let $\theta = (\theta_u, \theta_p)$ be the parameters of the payoff function and transition distribution. Let's suppose we have data on M markets each with N players and T_m time periods. We will assume that all of the data is generated by a single equilibrium. There may be multiple equilibria given the parameters of the model, but we only observe one.

4.2.1. *Maximum likelihood.* The likelihood can be written

$$\mathcal{L}(\theta, \mathrm{P}) = \sum_{m=1}^{M} \sum_{t=1}^{T_m} \sum_{i=1}^{N} \log \Lambda(a_{imt}|v_i^P(\cdot, x_{mt}; \theta)).$$

The maximum likelihood estimator can be written

$$\hat{\theta}_{MLE} = \arg\max_{\theta \in \Theta} \sup_{\mathrm{P} \in (0,1)^{N|X|}} \mathcal{L}(\theta, \mathrm{P}) \text{ s.t. } \mathrm{P} = \mathbf{\Lambda}(v^P(\theta))$$

This estimator is often very difficult to compute. In addition to the difficulty of repeatedly solving the dynamic programming problem for each player, we must find all equilibria for each θ at which we evaluate the constraint. In general, there multiple equilibria cannot be ruled out theoretically. There is also not any very easy method to compute all equilibria. As in the single agent case, we could perform two or three step maximum likelihood if there are some parameters that can be estimated in a first step without solving the full model.

4.2.2. *Pseudo likelihood.* As in the single agent case, we can also estimate θ using (nested) pseudo maximum likelihood. Given an initial consistent estimate of P, say $\hat{\mathrm{P}}$. The pseudo maximum likelihood estimate is

$$\hat{\theta}_{(0)}^{PMLE} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta, \hat{\mathrm{P}})$$

As above, we can repeat this process to get an iterated pseudo likelihood estimator. Let

$$\hat{\mathrm{P}}_{(1)} = \mathbf{\Lambda}\left(v^{\hat{\mathrm{P}}}(\hat{\theta}_{(0)}^{PMLE})\right).$$

Given that $\hat{P}$ is consistent, $\hat{\theta}_{(0)}^{PMLE}$ will be as well. Then $\hat{\mathrm{P}}_{(1)}$ is also consistent. We can then define

$$\hat{\theta}_{(1)}^{PMLE} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta, \hat{\mathrm{P}}_{(1)}).$$

Repeating this k times, we get the k-step pseudo maximum likelihood estimate, $\hat{\theta}_{(k)}^{PMLE}$. We can repeat this process to convergence. Unfortunately, this limit need not be the full maximum likelihood estimate. It only needs to solve

$$\hat{\theta}_{(\infty)}^{PMLE} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta, \hat{\mathrm{P}}_{(\infty)}) \text{ s.t. } \hat{\mathrm{P}}_{(\infty)} = \mathbf{\Lambda}(v^{\hat{\mathrm{P}}_{(\infty)}}(\hat{\theta}_{(\infty)}^{PMLE})).$$

To get the maximum likelihood estimate, we must take the maximum of all the limit points of iterated pseudo maximum likelihood estimates. Aguirregabiria and Mira (2007) call this the nested pseudo likelihood estimator.

4.2.3. *Asymptotic distribution of likelihood estimators.* Unlike in the single agent case, two-step maximum likelihood and pseudo maximum likelihood do not have the same asymptotic distribution. Pseudo maximum likelihood, iterated pseudo likelihood, and nested pseudo likelihood also have different distributions. To simplify the asymptotics, we will look at the case where T and N are xed, $M \to \infty$, and observations across markets are independent. The analysis can easily be adapted to other cases.

Lets begin by looking at the pseudo maximum likelihood estimator, $\hat{\theta}_{(0)}^{PLE}$. Taking the usual mean value expansion of the rst order condition, we have

$$\begin{aligned} 0 =& \nabla_\theta \mathcal{L}(\hat{\theta}_{(0)}^{PLE}, \hat{\mathrm{P}}) \\ =& \nabla_\theta \mathcal{L}(\theta_0, \mathrm{P}_0) + \nabla^2_{\theta,\theta}\mathcal{L}(\hat{\theta}_{(0)}^{PLE} - \theta_0) + \nabla^2_{\theta,P}\mathcal{L}(\hat{\mathrm{P}} - \mathrm{P}_0) + o_p(M^{-1/2}) \\ \sqrt{M}(\hat{\theta}_{(0)}^{PLE} - \theta_0) =& -\Omega_{\theta,\theta}\left(\sqrt{M}\nabla_\theta \mathcal{L} + \Omega_{\theta,P}\sqrt{M}(\hat{\mathrm{P}} - \mathrm{P}_0)\right) + o_P(1). \end{aligned}$$

Note that if we knew P_0, we could maximize the likelihood respect to θ given the known P to obtain an infeasible $\hat{\theta}^{IMLE}$ maximum liklehood estimate.

$$\hat{\theta}^{IMLE} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta, \mathrm{P}_0)$$

This infeasible estimate would have a similar asymptotic expansion as $\hat{\theta}_{(0)}^{PLE}$, but without the second term,

$$\sqrt{M}(\hat{\theta}^{IMLE} - \theta_0) = -\Omega_{\theta,\theta}\sqrt{M}\nabla_\theta \mathcal{L}.$$

Lets assume that

$$\sqrt{M}\nabla_\theta \mathcal{L} \xrightarrow{d} N(0, \Omega_{\theta,\theta})$$

Since $\hat{\theta}^{IMLE}$ is the maximum likelihood estimate of θ when P is known (and the constraint that P must correspond to equilibrium choice probabilities given θ has not been imposed), it is efcient (among estimates that do not impose the constraint). $\hat{\theta}_{(0)}^{PLE}$ is another consistent estimator. As usual, the difference between an efcient estimate and another estimate must be uncorrelated with the efcient estimate (If not, we could construct a more efcient estimate). Therefore, if we assume that

$$\sqrt{M}(\hat{\mathrm{P}} - \mathrm{P}_0) \xrightarrow{d} N(0, \Sigma_0)$$

then

$$\sqrt{M}(\hat{\theta}_{(0)}^{PLE} - \theta_0) \xrightarrow{d} N\left(0, \Omega_{\theta,\theta}^{-1} + \Omega_{\theta,\theta}^{-1}\Omega_{\theta,P}\Sigma\Omega_{\theta,P}'\Omega_{\theta,\theta}^{-1}\right)$$

From this, we see that $\hat{\theta}_{(0)}^{PLE}$ has variance equal to that of the infeasible MLE , $\Omega_{\theta,\theta}^{-1}$, plus an additional term due the estimation of P.

From the above, we can easily get an expression for the asymptotic distribution of iterated pseudo likelihood. We instantly get that

$$\sqrt{M}(\hat{\theta}_{(K)}^{PLE} - \theta_0) \xrightarrow{d} N\left(0, \Omega_{\theta,\theta}^{-1} + \Omega_{\theta,\theta}^{-1}\Omega_{\theta,P}\Sigma_k\Omega_{\theta,P}'\Omega_{\theta,\theta}^{-1}\right) \tag{5}$$

where Σ_k is the asymptotic variance of $\hat{\mathrm{P}}_k$. We can compute Σ_k using the delta method. For $k = 1$ we have

$$\begin{aligned} \sqrt{M}(\hat{\mathrm{P}}_{(1)} - \mathrm{P}_0) =& \sqrt{M}\left(\mathbf{\Lambda}(\hat{\theta}_{(0)}^{PLE}, \hat{P}_0) - \mathbf{\Lambda}(\theta_0, \mathrm{P}_0)\right) \\ =& \sqrt{M}\mathbf{\Lambda}_\theta(\hat{\theta}_{(0)}^{PLE} - \mathrm{P}_0) + \mathbf{\Lambda}_P(\hat{\theta}_{(0)}^{PLE} - \mathrm{P}_0) + o_P(1) \\ \xrightarrow{d}& N\left(0, \mathbf{\Lambda}_\theta\Omega_{\theta,\theta}^{-1}\mathbf{\Lambda}_\theta' + \left(\mathbf{\Lambda}_\theta\Omega_{\theta,\theta}^{-1}\Omega_{\theta,P} + \mathbf{\Lambda}_P\right)\Sigma_0\left(\mathbf{\Lambda}_\theta\Omega_{\theta,\theta}^{-1}\Omega_{\theta,P} + \mathbf{\Lambda}_P\right)'\right) \end{aligned}$$

We can plug this result into 5 to get the asymptotic variance of $\hat{\theta}^{PLE}_{(k)}$. However, there does not seem to be any nice way to simplify this expression. It is generally not even possible to show whether iterating decreases or increases the variance.

The nest pseudo likelihood estimate satisfies both $\Lambda(\hat{\theta}^{NPL}, \hat{\mathrm{P}}^{NPL}) = \mathrm{P}^{NPL}$ and the first order condition for maximizing the likelihood. Combining these two conditions and rearranging, we can show that

$$\sqrt{M}(\hat{\theta}^{NPL} - \theta_0) \xrightarrow{d} N\left(0, \left(\Omega_{\theta,\theta} + \Omega_{\theta,P}(I - \mathbf{\Lambda}_P)^{-1}\mathbf{\Lambda}_\theta\right)^{-1} \Omega_{\theta,\theta}\left(\Omega_{\theta,\theta} + \Omega_{\theta,P}(I - \mathbf{\Lambda}_P)^{-1}\mathbf{\Lambda}_\theta\right)^{-1'}\right)$$

Aguirregabiria and Mira (2007) show that if the eigenvalues of $\mathbf{\Lambda}_P$ are between 0 and 1, then this variance is smaller than $\Omega^{-1}_{\theta,\theta}$. However, there no particular reason for this condition to hold. **?** give some examples where the eigenvalues of Lb_P are outside of this range.

Comparison with single-agent model. In the single agent case, the pseudo likelihood estimate and iterated versions of it all have the same asymptotic distribution. The reason is the $\Omega_{\theta,P} = 0$ in single agent models. The reason this derivative is zero is that it depends on the derivative of v with respec to P. Since P maximizes v, this derivative is zero. In a game, v^P_i still has zero derivative with respect to P_i, but it doees not have zero derivative with respect to P_{-i}. Therefore, $\Omega_{\theta,P} \neq 0$ in dynamic games.

REFERENCES

AGUIRREGABIRIA, V., AND P. MIRA (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70(4), pp. 1519–1543.

——— (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75(1), pp. 1–53.

——— (2010): "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, 156(1), 38 – 67, ¡ce:title¿Structural Models of Optimization Behavior in Labor, Aging, and Health¡/ce:title¿.

BAJARI, P., V. CHERNOZHUKOV, H. HONG, AND D. NEKIPELOV (2009): "Nonparametric and Semiparametric Analysis of a Dynamic Discrete Game," Discussion paper.

BRENDT, E., R. HALL, B. HALL, AND J. HAUSMAN (1974): "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, pp. 653–665.

HOTZ, V. J., AND R. A. MILLER (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60(3), pp. 497–529.

MAGNAC, T., AND D. THESMAR (2002): "Identifying Dynamic Discrete Decision Processes," *Econometrica*, 70(2), pp. 801–816.

RUST, J. (1994): "Chapter 51 Structural estimation of markov decision processes," vol. 4 of *Handbook of Econometrics*, pp. 3081 – 3143. Elsevier.

SCHRIMPF, P. (2011): "Identification and estimation of dynamic games with continuous states and controls," Discussion paper, MIT.

STOKEY, N. L., AND R. E. LUCAS (1989): *Recursive methods in economic dynamics*. Harvard University Press.