Single-agent dynamic optimization models Part II Hotz Miller

C.Conlon - Adapted from M. Shum

Grad IO III

Spring 2012

Alternative approaches to estimation: avoid numberic dynamic programming

- One problem with Rust approach to estimating dynamic discrete-choice model very computer intensive. Requires using numerical dynamic programming to compute the value function(s) for every parameter vector θ.
- Alternative method of estimation, which avoids explicit DP. Present main ideas and motivation using a simplified version of Hotz and Miller (1993), Hotz, Miller, Sanders, and Smith (1994).
- For simplicity, think about Harold Zurcher model.

Back to Rust

- What do we observe in data from DDC framework? For bus i, time t, observe:
- $\{\tilde{x}_{it}, d_{it}\}$: observed state variables $\bar{x}_{it}$ and discrete decision (control) variables d_{it}. For simplicity, assume d_{it} is binary, $\in \{0, 1\}$
- Let $i = 1, \cdots, N$ index the buses, $t = 1, \cdots, T$ index the time periods.
- For Harold Zurcher model: $\tilde{x}_{it}$ is mileage on bus i in period t, and d_{it} is whether or not engine of bus i was replaced in period t.
- Given renewal assumptions (that engine, once repaired, is good as new), define transformed state variable x_{it}: mileage since last engine change.
- Unobserved state variables: $\epsilon_{it}, i.i.d.$ over i and t. Assume that distribution is known (Type 1 Extreme Value in Rust model)

Back to Rust

- In the following, let quantities with hats ($\hat{\cdot}$'s) denote objects obtained just from data.
- Objects with tildes ($\tilde{\cdot}$'s) denote "predicted" quantities, obtained from both data and calculated from model given parameter values θ.
- From this data alone, we can estimate (or "identify"):

$$\hat{G}(x'|x,d) \equiv \begin{cases} \sum_{i=1}^{N}\sum_{t=1}^{T-1} \frac{1}{\sum_i \sum_t 1(x_{it}=x, d_{it}=0)} \cdot \mathbf{1}(x_{i,t+1} \leq x', x_{it}=x, d_{it}=0), & \text{if } d=0 \\ \sum_{i=1}^{N}\sum_{t=1}^{T-1} \frac{1}{\sum_i \sum_t 1(d_{it}=1)} \cdot \mathbf{1}(x_{i,t+1} \leq x', d_{it}=1), & \text{if } d=1 \end{cases}$$

Alternative approaches to estimation

- Choice probabilities, conditional on state variable: $Pr(d=1|x)$, estimated by :

$$\hat{P}(d=1|x) \equiv \sum_{i=1}^{N}\sum_{t=1}^{T-1} \frac{1}{\sum_i \sum_t \mathbf{1}(x_{it}=x) \cdot \mathbf{1}(d_{it}=1, x_{it}=x).}$$

- Since $Pr(d=0|x) = 1 - Pr(d=1|x)$, we have $\hat{P}(d=0|x) = 1 - \hat{P}(d=1|x)$
- Let $\tilde{V}(x,d;\theta)$ denote the choice-specific value function, minus the error term ϵ_d. That is,

$$\tilde{V}(x,d;\theta) \equiv \tilde{V}(x,\epsilon,d;\theta) - \epsilon_d$$

By stationarity, note we do not index this probability explicitly with time t.

Alternative approaches to estimation

With estimates of $\hat{G}(\cdot|\cdot)$ and $\hat{p}(\cdot.\cdot)$, as well as a parameter vector θ, you can estimate the choice-specific value functions by constructing

$$
\begin{aligned}
\tilde{V}(x, d=1;\theta) &= u(x, d=1;\theta) + \beta E_{x'|x,d=1} E_{d'|x'} E_{\epsilon'|d',x'}[u(x', d';\theta) \\
+\epsilon' &+ \beta E_{x''|x',d'} E_{d''|x''} E_{\epsilon'|d'',x''}[u(x'', d'';\theta) + \epsilon'' + \beta \cdots]] \\
\tilde{V}(x, d=0;\theta) &= u(x, d=0;\theta) + \beta E_{x'|x,d=0} E_{d'|x'} E_{\epsilon'|d',x'}[u(x', d';\theta) \\
+\epsilon' &+ \beta E_{x''|x',d'} E_{d''|x''} E_{\epsilon'|d'',x''}[u(x'', d'';\theta) + \epsilon'' + \beta \cdots]]
\end{aligned}
$$

Here $u(x, d;\theta)$ denotes the per-period utility of taking choice d at state x, without the additive logit error. Note that the observation of $d'|x'$ is crucial to being able to forward-simulate the choice specific value functions. Otherwise, $d'|x'$ is multinomial with probabilities given below, and is impossible to calculate without knowledge of the choice-specific value functions.

Alternative approaches to estimation

In practice, "truncate" the infinite sum at some period T:

$$\tilde{V}(x, d = 1; \theta) =$$
$$u(x, d = 1; \theta) + \beta E_{x'|x,d=1} E_{d'|x'} E_{\epsilon''|d',x'} [u(x', d'; \theta) + \epsilon'$$
$$+\beta E_{x''|x',d''} E_{d''|x''} E_{\epsilon'|d'',x''} [u(x'', d''; \theta) + \epsilon'' + \cdots$$
$$\beta E_{x^T|x^{T-1},d^{T-1}} E_{d^T|x^T} E_{\epsilon^T|d^T,x^T} [u(x^T, d^T; \theta) + \epsilon^T]]]$$

Also, the expectation $E_{\epsilon|d,x}$ denotes the expectation of the ϵ conditional choice d being taken, and current mileage x. For the logit case, there is a closed form:

$$E[\epsilon|d, x] = \gamma - \log(PR(d|x))$$

where γ is Euler's constant $(0.577 \cdots)$ and $Pr(d|x)$ is the choice probability of action d at state x.

Both of the other expectations in the above expressions are observed directly from the data.

Alternative approaches to estimation

Choice-specific value functions can be simulated by (for $d = 1, 2$):

$$\begin{aligned}\tilde{V}(x,d;\theta) \approx = \quad & \tfrac{1}{S}\textstyle\sum_s [u(x,d;\theta) + \beta[u(x'^s, d'^s;\theta) + \gamma - \log(\hat{P}(d'^s|x'^s)) \\ & +\beta[u(x''^s, d''^s;\theta) + \gamma - \log(\hat{P}(d''^s|x''^s)) + \beta \cdots]]]\end{aligned}$$

where:

- $x'^s \sim \hat{G}(\cdot|x,d)$
- $d'^s \sim \hat{p}(\cdot|x'^s), x''^s \sim \hat{G}(\cdot|x'^s, d'^s)$
- & etc.

In short, you simulate $\tilde{V}(x,d;\theta)$ by drawing S sequences of (d_t, x_t) with a initial value of (d, x), and computing the present-discounted utility correspond to each sequence. Then the simulation estimate of $\tilde{V}(x,d;\theta)$ is obtained as the sample average.

Alternative approaches to estimation

- Given an estimate of $\tilde{V}(\cdot, d; \theta)$, you can get the *predicted choice probabilities*:

$$\tilde{p}(d=1|x;\theta) \equiv \frac{\exp\left(\tilde{V}(x, d=1;\theta)\right)}{\exp\left(\tilde{V}(x, d=1;\theta)\right) + \exp\left(\tilde{V}(x, d=0;\theta)\right)} \quad (1)$$

and analogously for $\tilde{p}(d=0|x;\theta)$. Note that the predicted choice probabilities are different from $\hat{p}(d|x)$, which are the actual choice probabilities computed from the actual data. The predicted choice probabilites depend on the parameters θ, whereas $\hat{p}(d|x)$ depend solely on the data.

Hotz Miller (1993) Estimator

- One way to estimate θ is to minimize the distance between the predicted conditional choice probabilities, and the actual conditional choice probabilities:

$$\hat{\theta} = \arg\min_{\theta} ||\hat{\mathbf{p}}(d=1|x) - \tilde{\mathbf{p}}(d=1|x;\theta)||$$

where **p** denotes a vector of probabilities, at various values of x. Another way to estimate θ is very similar to the Berry/BLP method. We can calculate directly from the data.

$$\hat{\delta}_x \equiv \log \hat{p}(d=1|x) - \log \hat{p}(d=0|x)$$

Alternative approaches to estimation

Given the logit assumption, from Eq. (1), we know that

$$\log \tilde{p}(d = 1|x) - \log \tilde{p}(d = 0|x) = \left[\tilde{V}(x, d = 1) - \tilde{V}(x, d = 0)\right].$$

Hence, by equation $\tilde{V}(x, d = 1) - \tilde{V}(x, d = 0)$ to $\hat{\delta}_x$, we obtain an alternative estimator for θ:

$$\bar{\theta} = \arg\min_{\theta} ||\hat{\delta}_x - \left[\tilde{V}(x, d = 1; \theta) - \tilde{V}(x, d = 0; \theta)\right]||.$$

Rust and Hotz-Miller Comparison

Rust's NFXP Algorithm

$$\begin{aligned} V_\theta(x) &= f(V_\theta(x), s, \theta) \Rightarrow f^{-1}(s,\theta) \\ P(d|x,\theta) &= g(V_\theta(x), s, \theta) \\ P(d|x,\theta) &= g(f^{-1}(s,\theta)) \end{aligned}$$

- At every guess of θ we solve the fixed point inverse
- Plug that in to get choice probabilities
- Evaluate the likelihood

Hotz-Miller (1993) to Aguirregabiria and Mira (2002)

- Choice probabilities conditional on any value of observed state variables are uniquely determined by the vector of normalized value functions
- HM show invertibility proposition (under some conditions).
- If mapping is one-to-one we can also express value function in terms of choice probabilities.

$$
\begin{aligned}
V_\theta(x) &= h(P(d|x,\theta), s, \theta) \\
P(d|x,\theta) &= g(V_\theta(x), s, \theta) \\
\Rightarrow P(d|x,\theta) &= g(h(P(d|x,\theta), s, \theta), s, \theta)
\end{aligned}
$$

- The above fixed point relation is used in Aguirregabiria and Mira (2002) in their NPL Estimation algorithm.

Hotz-Miller (1993) to Aguirregabiria and Mira (2002)

$$P^{k+1}(d|x,\theta) \quad = \quad g(h(\hat{P}^k(d|x,\theta),s,\theta),s,\theta)$$

- Key point here is that the functions $h(\cdot)$ and $g(\cdot)$ are quite easy to compute (compared to the inverse f^{-1}).
- We can substantially improve estimation speed by replacing P with $\hat{P}$ the Hotz-Miller simulated analogue.
- The idea is to reformulate the problem from value space to probability space.
- When initializing the algorithm with consistent nonparametric estimates of CCP, successive iterations return a sequence of estimators of the structural parameters
- Call this the K stage policy iteration (PI) estimator.

Hotz-Miller (1993) to Aguirregabiria and Mira (2002)

- This algorithm nests Hotz Miller $(K = 1)$ and Rust's NFXP $(K = \infty)$.
- Asymptotically everything has the same distribution, but finite sample performance may be increasing in K (at least in Monte Carlo).
- The Nested Pseudo Likelihood (NPL) estimator of AM $(K = 2)$ seems to have much of the gains.

A further shortcut in the discrete state case

- For the case when the state variable S are discrete, the value function is just a vector, and it turns out that, given knowledge of the CCP's $P(Y|S)$, solving the value function is just equivalent to solving a system of linear equations.
- This was pointed out in Pesendorfer and Schmidt-Dengler (2008) and Aguirregabiria and Mira (2007), and we follow the treatment in the latter paper

A further shortcut in the discrete state case

- Assume that choices Y and state variables S are all *discrete* (i.e. finite-valued).
- $|S|$ is cardinality of state space S. Here S includes just the observed state variables (not the unobserved shocks ϵ).
- Per-period utilities:

$$u(Y, S, \epsilon_Y; \Theta) = \bar{u}(Y, S; \Theta) + \epsilon_Y$$

- where ϵ_Y, for $y = 1, \cdots Y$, are i.i.d. extreme value distributed with unit variance.
- Parameters Θ. The discount rate β is treated as known and fixed.

A further shortcut in the discrete state case

- Introduce some more specific notation. Along the optimal dynamic path, at state S and optimal action Y, the continuation utility is

$$\bar{u}(Y,S) + \epsilon_Y + \beta \sum_{S'} P(S'|S,Y)V(S').$$

- where $V(\cdot)$ is the "Ex-ante" or "expected" Bellman equation (before ϵ observed, and hence before the action Y is chosen):

$$\begin{aligned} V(S) \quad = \quad & \sum_{Y} [P(Y|S)\{\bar{u}(Y,S) + E(\epsilon_Y|Y,S)\}] + \\ & \beta \underbrace{\sum_{Y} \sum_{S'} P(Y|S)P(S'|S,Y)V(S')}_{\sum_{S'} P(S'|S)V(S')} \end{aligned} \tag{2}$$

- (Similar to Rust's $EV(\cdots)$ function, but not the same.)

Matrix Version

$$\begin{aligned}\bar{V}(\Theta) &= \sum_{Y \in (0,1)} P(Y) * [\bar{u}(Y;\Theta) + \epsilon(Y)] + \beta \cdot F \cdot \bar{V}(\Theta) \\ \Leftrightarrow \bar{V}(\Theta) &= (I - \beta F)^{-1} \left\{ \sum_{Y \in (0,1)} P(Y) * [\bar{u}(Y;\Theta) + \epsilon(Y)] \right\}\end{aligned} \tag{3}$$

- $\bar{V}(\Theta)$ is the vector (each element denotes a differenet value of S for the expected value function at the parameter Θ
- '*' denotes elementwise multiplication
- F is the $|S|$-dimensional square matrix with (i, j)-element equal to $Pr(S' = j | S = i)$.
- $P(Y)$ is the $|S|$-vector consisting of elements $Pr(Y|S)$.
- $\bar{u}(Y)$ is the $|S|$-vector of per-period utilities $\bar{u}(Y; S)$
- $\epsilon(Y)$ is an $|S|$-vector where each element is $E[\epsilon_Y|Y, S]$. For the logit assumptions, the closed-form is

$$E[\epsilon_Y|Y, S] = 0.57721 - \log(P(Y|S)). \tag{4}$$

Pesendorfer Schmidt-Dengler (2008)

Based on this representation, P/S-D propose a class of "least-squares" estimators, which are similar to HM-type estimators, except now we don't need to "forward-simulate" the value function. For instance:

- Let $\hat{P}(\bar{Y})$ denote the estimated vector of conditional choice probabilities, and $\hat{F}$ be the estimated transition matrix. Both of these can be estimated directly from the data.
- For each posited parameter value Θ, and given $(\hat{F}, \hat{P}(\bar{Y}))$ use Eq. (3) to evaluate the value function $\bar{V}(X, \Theta)$, and derive the vector $P(\bar{Y}; \Theta)$ of implied choice probabilities at Θ, which has elements

$$P(Y|S;\Theta) = \frac{\exp[\bar{u}(Y,S;\Theta) + E_{S'|S,Y} V(S';\Theta)]}{\sum_Y \exp[\bar{u}(Y,S;\Theta) + E_{S'|S,Y} V(S';\Theta)]}.$$

- Hence, Θ can be estimated as the parameter value minimizing the norm $||\hat{P}(\bar{Y}) - P(Y;\Theta)||$.

Continuous state case

Recently, Srisuma and Linton (2009) have extended this approach to the case when the state variables S are continuous (but actions Y are still discrete). With continuous S, the state transition $p(S'|S)$ and choice probabilities $p(Y|S)$ for $Y \in (0, 1)$ are continuous functions. But they can still be estimated from the data (as discussed in the SL paper).

Continuous state case

Define $\bar{U}_\Theta(S) = \sum_Y P(Y|S) * [\bar{u}(Y,S;\Theta) + \epsilon(Y,S)]$, which can be calculated for each value of Θ. Now the ex-ante Bellman equation (2) becomes

$$V_\Theta(S) = \bar{U}_\Theta(S) + \beta \int p(S'|S) V_\Theta(S')$$

which can be expressed as the integral equation

$$\begin{aligned} V_\Theta &= \bar{U}_\Theta + \mathcal{L}_{\beta p(S'|S)} V_\Theta \\ &\Rightarrow (I - \mathcal{L}_{\beta p(S'|S)}) V_\Theta = \bar{U}_\Theta \end{aligned}$$

In the above, V_Θ and $\bar{U}_\Theta$ are viewed as elements in some space of functions of S, and $\mathcal{L}_{\beta p(S'|S)}$ denotes an integral operator, and I is an identity operator. The above is known as a *Friedholm* integral equation of the second kind.

Srisuma and Linton (2009)

Given that $p(S'|S)$ can be estimated from the data, and the function $\bar{U}_\Theta(S)$ can be estimated from the data for all S given a value of Θ, we can invert this equation to obtain an estimate of the value function:

$$\hat{V}_\Theta = (I - \mathcal{L}_{\beta\hat{p}(S'|S)})^{-1}\hat{\bar{U}}_\Theta.$$

This inversion can be done by geometric approximation:

$$\hat{V}_\Theta \approx (1 + \mathcal{L}_{\beta\hat{p}(S'|S)} + \mathcal{L}^2_{\beta\hat{p}(S'|S)} + \mathcal{L}^3_{\beta\hat{p}(S'|S)} + \cdots)\hat{\bar{U}}_\Theta.$$

Then this value function can be used in the estimation procedures in the previous sections (Note: $\mathcal{L}^2_{\beta\hat{p}(S'|S)}\hat{\bar{U}}_\Theta =$ $\int\int p(S''|S')p(S'|S)U(S'')dS''dS' = \int p(S''|S)U(S'')dS''$.)

(Note: if the above is simulated, then essentially you are back to forward simulation.)

Semiparametric Identification of DDC Models

We can also use the Hotz-Miller estimation scheme as a basis for an argument regarding the identification of the underlying DDC model. In Markovian DDC models, without unobserved state variables, the Hotz-Miller routine exploits the fact that the Markov probabilities $x', d'|x, d$ are identified directly from the data, which can be factorized into

$$f(x', d'|x, d) = \underbrace{f(d'|x')}_{\text{CCP}} \cdot \underbrace{f(x'|x, d)}_{\text{state law of motion}} \tag{5}$$

Semiparametric identification of DDC Models

In this section, we argue that once these "reduced form" components of the model are identified, the remaining parts of the modles – particularly, the per-period utility functions – can be identified without any further parametric assumptions. These arguments are drawn from Magnac and Thesmar (2002) and Bajari, Chernozhukov, Hong, and Nekipelov (2007).

We make the following assumptions, which are standard in this literature:

Semiparametric identification of DDC Models

1. Agents are optimizing in an infinite-horizon, stationary setting. Therefore, in the rest of this section, we use primes to denote the next-period values.
2. Actions D are chosen from the set $\mathcal{D} = \{0, 1, \cdots, K\}$.
3. The state variables are X.
4. The per-period utility from taking action $d \in \mathcal{D}$ in period t is:

$$u_d(X_t) + \epsilon_{d,t}, \forall d \in \mathcal{D}.$$

5. The $\epsilon_{d,t}$'s are utility shocks which are independent of X_t, and distributed i.i.d. with known distribution $F(\epsilon)$ across periods t and actions d. Let $\vec{\epsilon}_t \equiv (\epsilon_{0,1}, \epsilon_{1,t}, \cdots, \epsilon_{K,t})$.

Semiparametric identification of DDC Models

6. From the data, the conditional choice probabilities (CCP's)

$$p_d(X) \equiv \text{Prob}(D = 1|X)$$

   and the Markov transition kernel for X, denoted $p(X'|D, X)$, are identified.

7. $u_0(S)$, the per-period utility from $D = 0$, is normalized zero, for all X.

8. β, the discount factor, is known.

Magnac and Thesmar (2002) discuss the possibility of identifying β via exclusion restrictions, but we do not pursue that here.

Semiparametric identification of DDC Models

Following the arguments in Magnac and Thesmar (2002) and Bajari, Chernozhukov, Hong, and Nekipelov (2007), we will show the nonparametric identification of $u_d(\cdot)$, $d = 1, \cdots, K$, the per-period utility functions for all actions except $D = 0$.

The Bellman equation for this dynamic optimization problem is

$$V(X, \vec{\epsilon}) = \max_{d \in \mathcal{D}}(u_d(X) + \epsilon_d + \beta E_{X', \vec{\epsilon}|D,X} V(X', \vec{\epsilon}))$$

where $V(X, \vec{\epsilon})$ denotes the value function.

We define the choice-specific value function as

$$V_d(X) \equiv u_d(X) + \beta E_{X', \vec{\epsilon}|D,X} V(X', \vec{\epsilon}').$$

Semiparametric identification of DDC Models

Given these definitions, an agent's optimal choice when the state is X is given by

$$y^*(X) = \text{argmax}_{y \in \mathcal{D}}(V_d(X) + \epsilon_d)$$

Hotz and Miller (1993) and Magnac and Thesmar (2002) show that in this setting, there is a known one-to-one mapping, $q(X) : \mathbb{R}^K \to \mathbb{R}^K$, which maps the K-vector of choice probabilities $(p_1(X), \cdots, p_K(X))$ to the K-vector $(\triangle_1(X), \cdots, \triangle_K(X))$, where $\triangle_d(X)$ denotes the difference in choice-specific value functions

$$\triangle_d(X) \equiv V_d(X) - V_0(X)$$

Let the i-th element of $q(p_1(X), \cdots, p_K(X))$, denoted $q_i(X)$, be equal to $\triangle_i(X)$. The known mapping q derives just from $F(\epsilon)$, the known distribution of the utility shocks.

Semiparametric identification of DDC Models

Hence, since the choice probabilities can be identified from the data, and the mapping q is known, the value function differences $\triangle_1(X), \cdots, \triangle_K(X)$ is also known.

Next, we note that the choice-specific value function also satisfies a Bellman-like equation:

$$\begin{aligned} V_d(X) =& u_d(X) + \beta E_{X'|X,D}\left[E_{\vec{\epsilon}} \text{max}_{d' \in \mathcal{D}}(V_{d'}(X') + \epsilon'_{d'})\right] \\ =& u_d(X) + \beta E_{X'|X,D}\left\{V_0(X') + \left[E_{\vec{\epsilon}} \text{max}_{d' \in \mathcal{D}}(\triangle_{d'}(X') + \epsilon'_{d'})\right]\right\} \\ =& u_d(X) + \beta E_{X'|X,D}[H(\triangle_1(X'), \cdots, \triangle_K(X')) + V_0(X')] \end{aligned} \tag{6}$$

where $H(\cdots)$ denotes McFadden's "social surplus" function, for random utility models (cf. Rust (1994, pp. 3104ff)). Like the q mapping, H is a known function, which depends just on $F(\epsilon)$, the known distribution of the utility shocks.

Semiparametric identification of DDC Models

Using the assumption that $u_0(X) = 0$, $\forall X$, the Bellman equation for $V_0(X)$ is

$$V_0(X) = \beta E_{X'|X,D}[H(\triangle_1(X'), \cdots, \triangle_K(X')) + V_0(X')] \qquad (7)$$

In this equation, everything is known (including, importantly, the distribution of $X'|X,D$), except the $V_0(\cdot)$ function. Hence, by iterating over Eq. (6), we can recover the $V_0(X)$ function. Once $V_0(\cdot)$ is known, the other choice-specific value functions can be recovered as

$$V_d(X) = \triangle_d(X) + V_0(X), \;\; \forall y \in \mathcal{D}, \forall X.$$

Semiparametric identification of DDC Models

Finally, the per-period utility functions $u_d(X)$ can be recovered from the choice-specific value functions as

$$u_d(X) = V_d(X) - \beta E_{X'|X,D}[H(\triangle_1(X'), \cdots, \triangle_K(X')) + V_0(X')], \forall y \in \mathcal{D}, \forall X$$

where everything on the right-hand side is known.

Semiparametric identification of DDC Models

REMARK: For the case where $F(\epsilon)$ is the Type 1 Extreme Value distribution, the social surplus function is

$$H(\triangle_1(X), \cdots, \triangle_K(X)) = \log\left[1 + \sum_{d=1}^{K} \mathsf{exp}(\triangle_d(X))\right]$$

and the mapping q is such that

$$q_d(X) = \triangle_d(X) = \log(p_d(X)) - \log(p_0(X)), \forall d = 1, \cdots, K,$$

where $p_0(X) \equiv 1 - \sum_{d=1}^{K} p_d(X)$.

References

AGUIRREGABIRIA, V., AND P. MIRA (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75, 1–53.

BAJARI, P., V. CHERNOZHUKOV, H. HONG, AND D. NEKIPELOV (2007): "Nonparametric and Semiparametric Analysis of a Dynamic Game Model," Manuscript, University of Minnesota

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

HOTZ, J., R. MILLER, S. SANDERS, AND J. SMITH (1994): "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, 61, 265–289.

References

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

PESENDORFER, M., AND P. SCHMIDT-DENGLER (2008): "Asymptotic Least Squares Estimators for Dynamic Games," *Review of Economic Studies*, 75, 901–928.

RUST, J. (1994): "Structural Estimation of Markov Decision Processes," in *Handbook of Econometrics, Vol. 4*, ed. by R. Engle, and D. McFadden, pp. 3082-146. North Holland.

SRISUMA, S., AND O. LINTON (2009): "Semiparametric Estimation of Markov Decision Processes with Continous State Space," manuscript, LSE.