Single-Agent Dynamic Discrete Choice

Ron Yang

March 5, 2019

This note draws on prior slides by Robin Lee and Myrto Kalouptsidi. I focus on (1) techniques you need for the problem set, and (2) material useful for building toward dynamic games.

1 Static Discrete Choice

1.1 Model Setup

I begin with the static discrete choice problem, where the payoff to choosing an action is the sum of a mean utility $\pi(a,x)$ and an action-specific random shock ϵ^a.

$$u(x,\epsilon) = \max_{a \in A} \pi(a,x) + \epsilon^a$$

Given a distributional assumption on $\epsilon^a \sim_{iid} G$, we can form choice probabilities.

$$P(a \mid x) = P(\pi(a,x) + \epsilon^a \geq \pi(a',x) + \epsilon^{a'}, \quad \forall a' \in A)$$

When ϵ^a is distributed i.i.d. logit, these choice probabilities are

$$p_a(x) = P(a \mid x) = \frac{\exp(\pi(a,x))}{\sum_{a' \in A} \exp(\pi(a',x))}$$

Let $p(x)$ denote a vector which stacks $p_a(x)$ for all a.

1.2 CCPs

The choice probabilities imply that, conditional on x, the higher the mean utility from taking an action, the more likely we are to observe that action being taken. This suggests that can we learn about conditional differences in mean utility across actions from differences in conditional choice probabilities.

This idea is formalized in the following theorem.

Theorem 1. *(Hotz-Miller Inversion, Static)*
$\forall (a,x) \in A \times X$, $j \in A$,

$$\pi(a,x) - \pi(j,x) = \phi_{aj}(p(x))$$

where $\phi_{aj} : \mathbb{R}^{|A|} \to \mathbb{R}$ only depends on G.

When G is logit, this theorem states that mean utility differences are equal to log probability differences.

$$\pi(a,x) - \pi(j,x) = \log p_a(x) - \log p_j(x)$$

If we fix a reference action a, this inversion allows us to express the agent's expected payoff conditional on x as a function of (1) the reference action's mean utility and (2) the choice probabilities.[1]

This idea is formalized in the following lemma.

Theorem 2. *(Arcidiacono-Miller Lemma, Static)*
$\forall (a,x) \in A \times X$, $\exists \psi_a(p(x)) : \mathbb{R}^{|A|} \to \mathbb{R}$ *s.t.*

$$\mathcal{E}[u(x,\epsilon) \mid x] = \pi(a,x) + \psi_a(p(x))$$

where ψ_a only depends on G.

When G is logit, this theorem states that expected payoff conditional on x is the reference action's mean utility less the log probability of choosing the reference action, plus Euler's constant γ[2].

$$\mathcal{E}[u(x,\epsilon) \mid x] = \pi(a,x) + \gamma - \log p_a(x)$$

Conceptually, our discrete choice model implies a mapping from payoffs $\pi(a,x)$ to choice probabilities $p_a(x)$. The Hotz-Miller Inversion and Arciadiacono-Miller Lemma say we can "invert" this mapping to recover mean utilities from observed choice probabilities. The intuition is straightforward: if agents choose a more often than b, the utility from a must be higher than b.

1.3 Identification

Let's use the inversion to understand what we can identify in the static discrete choice model. For simplicity, suppose that x is discrete with $|X|$ states.

We observe $|A| \times |X|$ CCP's $p_a(x)$, and we are interested in the $|A| \times |X|$ mean utilities $\pi(a,x)$. For each state x, fixing a reference action j, the Hotz-Miller Inversion gives you $|A| - 1$ equations of the form,

$$\forall a \neq j, \mid \pi(a,x) = \phi_{aj}(p(x)) + \pi(j,x)$$

To identify the conditional payoffs, we need $|X|$ additional restrictions. These additional restrictions might take the form of:

- A parametric assumption on $\pi(a,x)$.
- A state-independence assumption on the reference action's payoff $\pi(j,x) = \pi(j)$.

For example, in a static consumer choice setting, a natural restriction would be normalizing the outside option's mean utility, $\pi(\text{No Purchase} \mid x) = 0$. Notice that we are not normalizing the outside option's payoff; the outside option still receives a ϵ^a shock.

[1] Derivation:

$$\begin{aligned}
\mathcal{E}[u(x,\epsilon) \mid x] &= \int \max_{j \in A} \{\pi(j,x) + \epsilon^j\} dG(\epsilon) \\
&= \int \max_{j \in A} \{\pi(j,x) - \pi(a,x) + \epsilon^j\} dG(\epsilon) + \pi(a,x) \\
&= \int \max_{j \in A} \{\phi_{aj}(p(x)) + \epsilon^j\} dG(\epsilon) + \pi(a,x) \\
&= \phi_a(p(x)) + \pi(a,x)
\end{aligned}$$

[2] Derivation:

$$\begin{aligned}
\mathcal{E}[u(x,\epsilon)|x] &= \sum_a P(a|x)(\pi(a,x) + E[\epsilon^a|a,x]) \\
&= \sum_a P(a|x)(\pi(a,x) + \gamma - \log p_a(x)) \\
&= \sum_a P(a|x)(\pi(a,x) - \pi(j,x) + \gamma - \log p_a(x)) + \pi(j,x) \\
&= \sum_a P(a|x)(\log p_a(x) - \log p_j(x) + \gamma - \log p_a(x)) + \pi(j,x) \\
&= \pi(j,x) + \gamma - \log p_j(x)
\end{aligned}$$

2 Dynamic Discrete Choice

My notation will deviate from Robin's slides, but they are both internally consistent. I hope the second treatment will be clarifying.

2.1 Model Setup

- Let time $t \in \{0, 1, \dots, \infty\}$ be discrete. Let agents be indexed by i.
- For each time t, agent i has a vector of **state** variables $s_{it} = (x_{it}, \epsilon_{it})$ where x_{it} are observable and ϵ_{it} are unobservable.
- For each time t, agent i chooses an **action** $a_{it} \in A$.
- Given the current action and state, $F(s_{i,t+1} \mid \{a_{it}, s_{it}\})$ is the distribution of next-period state **transitions**. $F_{it}(s_{i,t+1} \mid \{a_{it}, s_{it}\})$ is agent i's perception of the distribution of next-period state transition at time t.
- The agent receives flow **payoffs** each period of $\pi(a_{it}, s_{it})$ which depend on the current action and state. Agents discount future payoffs by β.
- In summary, the **timing** of each period is: (1) agents observe s_{it}, (2) agents choose a_{it}, (3) agents receive payoff $\pi(a_{it}, s_{it})$, and (4) the next period $s_{i,t+1}$ is drawn.

The agent chooses a sequence of actions $\{a_{it}\}$ to maximize expected payoffs,

$$V(s_{i0}) = \max_{\{a_{it}\}} \mathcal{E}[\sum_{t=0}^{\infty} \beta^t (\pi(a_{it}, s_{it})) \mid a_{i0}, s_{i0}]$$

2.2 Assumptions

Rational Expectations: Agent perceptions of transition probabilities are correct.

$$F_{it}(s_{i,t+1} \mid \{a_{it}, s_{it}\}) = F(s_{i,t+1} \mid \{a_{it}, s_{it}\})$$

Markov: State transitions only depend on the current state and action.

$$F(s_{i,t+1} \mid \{a_{it}, s_{it}\}) = F(s_{i,t+1} \mid s_{it}, a_{it})$$

Conditional Independence: Unobservables are independent of observable states and actions. Conditional on current states and actions, observables are independent of unobservables.

$$F(s_{i,t+1} \mid s_{it}, a_{it}) = F(x_{i,t+1} \mid x_{it}, a_{it}) G(\epsilon_{i,t+1})$$

Additive Separability: The effect of unobservables on flow payoffs is additively separable. $\epsilon_{it} \in \mathbb{R}^A$ takes on one value ϵ_{it}^a for each action a. Per-period payoffs are the sum of a mean utility and a shock.

$$\pi(a_{it}, s_{it}) = \pi(a_{it}, x_{it}) + \epsilon_{it}^{a_{it}}$$

2.3 Value Functions

Given these assumptions, we can rewrite the agent's problem as a Bellman equation:

$$V(x_{it}, \epsilon_{it}) = \max_{a \in A} \{\pi(a, x_{it}) + \epsilon_{it}^a + \beta \mathcal{E}[V(x_{i,t+1}, \epsilon_{i,t+1}) \mid a, x_{it}]\}$$

The agent's **Value Function** $V(x_{it,}\epsilon_{it})$ is the expected discounted flow of payoffs conditional on the current state, including both observables and unobservables.

Since the value function is a function of observables, it is by construction difficult to work with. We define two useful alternative value functions;

The **Ex-Ante Value Function** $V(x_{it})$ is the value function conditional on current observables (but not current unobservables). Due to Conditional Independence, we can think of drawing the observables "first" and drawing the unobservables "second."

$$V(x_{it}) = \mathcal{E}[V(x_{it}, \epsilon_{it})] = \int V(x_{it}, \epsilon) dG(\epsilon)$$

The **Conditional Value Function** $v_a(x_{it})$ is the value function conditional on current observables and taking action a, less the effect of current unobservables.

$$v_a(x_{it}) = \pi(a, x_{it}) + \beta \mathcal{E}[V(x_{i,t+1}, \epsilon_{i,t+1}) \mid a, x_{it}]$$

In the static setting, $\pi(a, x)$ is the counterpart to the conditional value function; it summarizes the observable component of the utility from each action. The difference between the static and dynamic setting is the continuation value,

$$v_a(x_{it}) - \pi(a, x) = \beta \mathcal{E}[V(x_{i,t+1}, \epsilon_{i,t+1}) \mid a, x_{it}]$$

The three value functions are closely linked. The conditional value function summarizes the "observable" component of the value from each action. Adding the realization of current unobservables and taking a max yields the value function.

$$V(x_{it}, \epsilon_{it}) = \max_{a \in A} \{v_a(x_{it}) + \epsilon_{it}^a\}$$

Taking an expectation over current unobservables then yields the ex ante value function.

$$V(x_{it}) = \int \max_{a \in A} \{v_a(x_{it}) + \epsilon^a\} dG(\epsilon)$$

2.4 CCPs

Given these three value functions, we can now define the full dynamic versions of our two theorems.

Theorem 3. *(Hotz-Miller Inversion)*
$\forall (a, x) \in A \times X,\ j \in A,$

$$v_a(x) - v_j(x) = \phi_{aj}(p(x))$$

where $\phi_{aj} : \mathbb{R}^{|A|} \to \mathbb{R}$ *only depends on* G*.*

Theorem 4. *(Arcidiacono-Miller Lemma)*
$\forall (a, x) \in A \times X,\ \exists \psi_a(p(x)) : \mathbb{R}^{|A|} \to \mathbb{R}$ *s.t.*

$$V(x) = v_a(x) + \psi_a(p(x))$$

where ψ_a *only depends on* G*.*

The static intuition carries through: we can recover differences in conditional value functions from differences in conditional choice probabilities.

When G is logit, the difference in conditional value functions is the difference in log conditional choice probabilities,

$$v_a(x) - v_j(x) = \phi(p(x)) = \log p_a(x) - \log p_j(x)$$

When G is logit, the ex-ante value function is the conditional value function less the log conditional choice probability, plus Euler's constant.

$$V(x) = v_a(x) + \psi_a(p(x)) = v_a(x) + \gamma - \log p_a(x)$$

2.5 Identification

Setup Again, suppose that x is discrete with $|X|$ states, and suppose we know the discount factor β and the distribution of unobservables G. For ease of notation, let $F_a(x)$ be the vector of transition probabilities $F(x' \mid x = x, a = a)$, $\pi_a(x) = \pi(a, x)$, $\psi_a(x) = \psi_a(p(x))$, and $\phi_{aj}(x) = \phi_{aj}(p(x))$. When I omit the the state x, I am stacking across states to form a vector.

The econometrician observes the $|A| \times |X|$ conditional choice probabilities $p(x)$ and the $|A| \times |X| \times |X|$ transition probabilities $F_a(x)$, and wants to recover the $|A| \times |X|$ mean utilities $\pi_a(x)$, the $|X|$ ex ante values $V(x)$, and the $|A| \times |X|$ conditional values functions $v_a(x)$.

The conditional value functions give us $|A| \times |X|$ restrictions.

$$v_a(x) = \pi_a(x) + \beta F_a(x)V$$

The Hotz-Miller inversion gives us $|A - 1| \times |X|$ restrictions.

$$v_a(x) - v_j(x) = \phi_{aj}(x)$$

Arcidiacono-Miller Lemma gives us $|A| \times |X|$ restrictions.

$$V(x) = v_a(x) + \psi_a(x)$$

We are left needing $|X|$ additional restrictions to identify our objects of interest.

An Affine Formula Precisely, the above restrictions imply that all mean utilities are affine transformations of the mean utility of a reference action j.[3]

$$\pi_a = A\pi_j + [A\psi_j - \psi_a] \tag{1}$$

where $A = (I - \beta F_a)(I - \beta F_j)^{-1}$ captures the difference between doing a forever and doing j forever on the distribution of states visited.

Equation 1 tells us that in a DDC model, flow payoffs π_a are identified by difference in transition probabilities A and differences in CCPs ψ_a.

Up to the choice of π_j ($|X|$ restrictions), all other π_a's are known. As in the static case, we might make parametric assumptions on $\pi_a(x)$, or we might make assumptions about the state-independence of $\pi_j(x)$. Once π_a is known, the ex ante value function can be found by combining the conditional value function definition and the Arcidiacono-Miller Lemma.

$$V = (I - \beta F_a)^{-1}(\pi_a + \psi_a) \tag{2}$$

Intuition What does Equation 1 mean? Let's take this step by step.

$A\pi_j$ is the vector of flow payoffs such that the infinite discounted value of doing a forever with payoffs $A\pi_j$ is equal to the infinite discounted value of doing j with payoffs π_j. Naturally, if $F_a = F_j$, then $A\pi_j = \pi_j$.

$$A\pi_j = (I - \beta F_a)(I - \beta F_j)^{-1}\pi_j \iff (I - \beta F_a)^{-1}A\pi_j = (I - \beta F_j)^{-1}\pi_j$$

[3]Derivation: Stack the $|X|$ conditional value functions, and substitute in the Arcidiacono-Miller conditions. Mean utilities are the sum of

$$\begin{aligned}\pi_a(x) &= v_a(x) - \beta F_a(x)V \\ &= V - \psi_a(x) - \beta F_a(x)V \\ &= (I - \beta F_a)V - \psi_a(x)\end{aligned}$$

Fix a reference action j and invert to express the ex ante value function as a discounted sum of future payoffs and ψ to a reference action j.

$$V = (I - \beta F_j)^{-1}(\pi_j + \psi_j(x))$$

Plug this back into the mean utility.

Fixing π_j, variation in $A\pi_j$ reflects variation in the states that a policy of doing $a$will visit compared to the reference policy. If choosing a in state x is more likely to bring the agent to high payoff states than choosing j, then $(A\pi_j)(x)$ will be lower than π_j.

For more intuition about the second term, consider the logit case. From above, ψ_a is Euler's constant less the log CCP. This implies, for $\mathbf{1}$ = a vector of one's,

$$\begin{aligned} A\psi_j - \psi_a &= A(\gamma\mathbf{1} - \log p_j) - (\gamma\mathbf{1} - \log p_a) \\ &= A \log \frac{p_a}{p_j} - (A - I)\gamma\mathbf{1} \end{aligned}$$

$A \log \frac{p_a}{p_j}$ is the per-period log-odds that would equalize the total discounted log-odds from doing a forever and doing j forever with per-period log-odds $\log \frac{p_a}{p_j}$.

$$(I - \beta F_a)^{-1} A \log \frac{p_a}{p_j} = (I - \beta F_j)^{-1} \log \frac{p_a}{p_j}$$

Variation in $A \log \frac{p_a}{p_j}$ reflects variation across actions in (1) CCPs and (2) future states visited. If a has a higher choice probability than j, then $\log \frac{p_a}{p_j}$ will be higher. If choosing a is more likely to bring the agent to states with a high probability of choosing a, then $A \log \frac{p_a}{p_j}$ will be lower than $\log \frac{p_a}{p_j}$.

Finally, $(A - I)\gamma\mathbf{1}$ reflects the expected value of the logit shocks; in this case, $(A - I)\gamma\mathbf{1} = \mathbf{0}$ so I ignore it[4].

Putting these intuitions together, what observable variation identifies π_a in a DDC model? A DDC model will predict that the flow payoff of an action must be high if (1) the action brings the agent to high-payoff states, (2) the action has a high choice probability, and (3) the action brings the agent to low-payoff states despite having a high choice probability.

3 Estimation

3.1 Setup

Goal We make some additional assumptions:

- To addressing the identification issues, assume a parametric form for the flow payoffs, $\pi_a(x \mid \theta)$ and the transition probabilities $F_a(x' \mid x, \theta)$. Let the true DGP be parameterized by θ_0.
- Assume we know β and G with certainty, where $\epsilon \sim G$ is i.i.d. logit.

We observe $\{a_{it}, x_{it}\}$ for agents i and time periods t, and are interested in estimating θ.

CCPs Given a guess for θ, our model will yield a CCP

$$P(a_t | x_t, \theta) = \frac{\exp(v_{a_t}(x_t \mid \theta))}{\sum_{j \in A} \exp(\cdot)} = \frac{\exp(\pi_{a_t}(x_t \mid \theta) + \beta\mathcal{E}[V(x_{t+1} \mid \theta) \mid x_t, a_t])}{\sum_{j \in A} \exp(\cdot)}$$

If we directly observed ex-ante value functions or conditional value functions, then we could just plug them in to get predicted CCPs. Given predicted CCPs, we could either use MLE and find θ to maximize the likelihood of observed choices, or we could use CMD and find θ to minimize the distance between observed and predicted CCPs.

Since we typically do not observe $V(x')$, our dynamic estimation techniques will be focused on estimating $V(x')$. Put another way, if we tried to do static discrete choice estimation when the structural model was dynamic, the omitted variable would be the expected ex-ante value functions.

[4]Derivation:

$$\begin{aligned} (A - I)\gamma\mathbf{1} &= ((I - \beta F_a)(I - \beta F_j)^{-1} - I)\gamma\mathbf{1} \\ &= \beta\gamma(F_j - F_a)(I - \beta F_j)^{-1}\mathbf{1} \end{aligned}$$

Each row of $(I - \beta F_j)^{-1}$ sums to β^{-1}, and each row of $F_j - F_a$ sums to 0, so $(A - I)\gamma\mathbf{1} = \mathbf{0}$. Intuitively, if the payoff is constant in every state, then it doesn't matter what the state transitions look like, $A\mathbf{1} = I\mathbf{1}$.

3.2 Nested Fixed Point Estimation (Rust)

For a given guess of θ, $V(x)$ is the solution to a standard dynamic programming problem.

$$V(x \mid \theta) = \mathcal{E}_\epsilon[V(x, \epsilon \mid \theta)] = \mathcal{E}_\epsilon[\max_{a \in A}\{\pi_a(x \mid \theta) + \epsilon^a + \beta \mathcal{E}_{x', \epsilon'}[V(x', \epsilon' \mid \theta) \mid a, x]\}]$$

When ϵ is logit,

$$V(x \mid \theta) = \log \sum_a \exp(\pi_a(x \mid \theta) + \beta \mathcal{E}_{x', \epsilon'}[V(x', \epsilon' \mid \theta) \mid a, x]) + \gamma$$

For discrete states, we can stack the value function into a convenient vector representation.

$$V(\theta) = \log \sum_a \exp(\pi_a(\theta) + \beta F_a(\theta) V(\theta) + \gamma$$

This problem can be solved for an optimal value function $V^*(x \mid \theta)$ using standard algorithms (value function iteration, policy function iteration, etc.) and an approximation to the value function (discretization, splines, polynomials, etc.). Given a solution, $V^*(x \mid \theta)$, we can proceed with the MLE we specified above. This is “Nested Fixed Point Estimation” because we nest a fixed-point computation within each evaluation of our likelihood for a given θ.

3.3 Conditional Choice Probability Estimation (Hotz-Miller)

Setup The NFP estimator relies on the DDC problem having a unique solution in the single-agent case. When we move to dynamic games with multiple equilibria, we can no longer guarantee that the value functions we compute match the value functions in the data.

Our discussion in (2.5) suggests an alternative: we can infer the equilibrium value function from observed CCPs. This will require a two-step estimator.

For convenience, I denote values estimated in the first stage by hats, and values estimated in the second stage by stars. Be careful of the difference between the true parameters θ_0 and our guessed parameters θ.

First Step In the first step, we estimate the CCPs $\hat{p}(\theta_0)$ and state transition probabilities $\hat{F}_a(\theta_0)$. Since we observe (x_t, a_t), any sufficiently flexible non-parametric estimator can work.

Using these CCPs, we can recover the conditional value functions (up to a normalization).

$$\hat{v}_a(\theta_0) - \hat{v}_j(\theta_0) = \phi_{aj}(\hat{p}(\theta_0))$$

Second Step In the second step, we use our first stage estimates to back out value functions.

These are a few options depending on the exact problem:

Finite State Space Case: Inversion (Hotz-Miller) When the state space is finite, we can use Equations 1 and 2 to derive an expression for the ex ante value function given a guess θ.

$$V^*(\theta) = (I - \beta \hat{F}_a(\theta_0))^{-1}(\pi_a(\theta) + \psi_a(\hat{p}(\theta_0)))$$

General Case: Simulation (Hotz-Miller-Sanders-Smith) The inversion method faces two limitations: it requires finite state spaces, and it requires a matrix inversion. When the state space is rich, requiring either a large number of discrete states or a richer approximation to the state space, the inversion method is infeasible.

For general state spaces, we can leverage the face that our estimated CCPs $\hat{p}(a \mid x, \theta_0)$ already reflect the optimal policy function for the true parameters θ_0 and the true equilibrium. We will take these CCPs as given and compute the values of v_a implied by our model and our guess of θ. If $\theta = \theta_0$, the true CCPs and model-implied CCPs will match.

The conditional value function can be expressed as a sum of expected flow payoffs over sequences of states $\{x_t\}$, where state transitions are driven by optimal behavior.

$$v_a(x \mid \theta) = \mathcal{E}[\sum_{t=0}^{\infty} \beta^t [\pi_{a_t}(x_t \mid \theta) + \epsilon_t^{a_t}]]$$
$$\text{s.t.} \quad a_t = \arg\max_{a \in A} \hat{v}_a(x_t \mid \theta_0) + \epsilon_t^a$$

We can use simulation to compute this expectation. Let S be the number of simulation draws and T be the number of simulation periods. To estimate $v_a^*(x \mid \theta)$ by simulation,

1. For each simulation draw $s \in S$, draw a sequence of state-action pairs $\{x_t, a_t\}_s$
   (a) Initialize the initial state at $x_0 = x$.
   (b) For each simulation period $t \in [0, \ldots, \infty]$,
      i. Choose $a_t = \arg\max_{a \in A} \hat{v}_a(x_t \mid \theta_0) + \epsilon_t^a$. Equivalently, draw $a_t \sim \hat{p}_a(x_t \mid \theta_0)$
      ii. Draw $x_{t+1} \sim \hat{F}(x_{t+1}, x_t, a_t \mid \theta_0)$.
      iii. Draw $\epsilon_{t+1} \sim Logit$.
2. For each simulated sequence of state-action pairs $\{x_t, a_t\}_s$, compute the conditional value function.
$$v_a^s(x \mid \{x_t, a_t\}_s, \theta) = \sum_{t=0}^{\infty} \beta^t [\pi_{a_t}(x_t \mid \theta) + \mathcal{E}[\epsilon_t^{a_t} \mid a_t, x_t]]$$
3. Average over simulation draws to get an estimate of the conditional value function.
$$v_a^*(x \mid \theta) = \frac{1}{S} \sum_s v_a^s(x \mid \{x_t, a_t\}_s, \theta)$$

Given conditional value functions, $v_a^*(x \mid \theta)$, we can compute predicted CCPs and proceed with estimation.

Several features can dramatically improve the speed of the simulation:

- The simulation paths $\{x_t, a_t\}_s$ only depend on your estimated probabilities and state transitions, so they can be held constant across guesses of θ.
- When ϵ_t is logit, there is an analytical form for $\mathcal{E}[\epsilon_t^{a_t} \mid a_t, x_t] = \gamma - \log \hat{p}_a(x)$.
- For some parameterizations of $\pi(a_t, x_t, \theta)$, $v_a^*(x \mid \theta)$ can be expressed as a linear function of θ.

4 Tips

1. When picking your first stage estimator, be mindful of the tradeoff between flexibility and sparsity. What are your predicted CCPs in states that have never been observed (but a simulated agent might enter)?
2. Don't forget the γ terms.
3. The Rust data is messy. If you make any data cleaning decisions, document and justify them.

5 References

- Arcidiacono, P., & Miller, R. A. (2011). Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity. Econometrica, 79(6), 1823–1867.
- Hotz, V. J., & Miller, R. A. (2006). Conditional Choice Probabilities and the Estimation of Dynamic Models. The Review of Economic Studies, 60(3), 497.

- Hotz, V. J., Miller, R. A., Sanders, S., & Smith, J. (2006). A Simulation Estimator for Dynamic Models of Discrete Choice. The Review of Economic Studies, 61(2), 265–289.
- Kalouptsidi, M., Scott, P. T., & Souza-Rodrigues, E. (2017). On the non-identification of counterfactuals in dynamic discrete games. International Journal of Industrial Organization, 50, 362–371.
- Rust, J. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of (Vol. 55).
- Rust, J. (1994). Structural Estimation of Markov Decision Processes. Handbook of Econometrics, 1.
- Thesmar, D., & Magnac, T. (2008). Identifying Dynamic Discrete Decision Processes. Econometrica, 70(2), 801–816.