Conditional Choice Probabilities
First of two lectures for:
DSE 2021 Summer School in Bonn

Robert A. Miller

Carnegie Mellon University

August 2021

A Class of Dynamic Discrete Choice Markov Models

Discrete time and finite choice sets

- Let $T \in \{1, 2, \ldots\}$ with $T \leq \infty$ denote the horizon of the optimization problem and $t \in \{1, \ldots, T\}$ denote the time period.
- Each period the individual chooses amongst J mutually exclusive actions.
- Let $d_t \equiv (d_{1t}, \ldots, d_{Jt})$ where $d_{jt} = 1$ if action $j \in \{1, \ldots, J\}$ is taken at time t and $d_{jt} = 0$ if action j is not taken at t.
- $x_t \in \{1, \ldots, X\}$ for some finite positive integer X for each t.
- $\epsilon_t \equiv (\epsilon_{1t}, \ldots, \epsilon_{Jt})$ where $\epsilon_{jt} \in \mathbb{R}$ for all (j, t).
- Assume the data comprises observations on (d_t, x_t).
- The joint mixed density function for the state in period $t+1$ conditional on (x_t, ϵ_t), denoted by $g_{t,x,\epsilon}(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t)$, satisfies the conditional independence assumption:

$$g_{t,j,x,\epsilon}(x_{t+1}, \epsilon_{t+1} | x_t, \epsilon_t) = g_{t+1}(\epsilon_{t+1} | x_{t+1}) f_{jt}(x_{t+1} | x_t)$$

where $g_t(\epsilon_t | x_t)$ is a conditional density for the disturbances, and $f_{jt}(x_{t+1} | x)$ is a transition probability for x conditional on (j, t).

A Class of Dynamic Discrete Choice Markov Models

Bounded additively separable preferences

- Denote the discount factor by $\beta \in (0,1)$ and the current payoff from taking action j at t given (x_t, ϵ_t) by $u_{jt}(x_t) + \epsilon_{jt}$.
- To ensure a transversality condition is satisfied, assume $\{u_{jt}(x)\}_{t=1}^{T}$ is a bounded sequence for each $(j, x) \in \{1, \ldots, J\} \times \{1, \ldots, X\}$, and so is:

$$\left\{ \int \max\left\{ |\epsilon_{1t}|, \ldots, |\epsilon_{Jt}| \right\} g_t\left(\epsilon_t | x_t\right) d\epsilon_t \right\}_{t=1}^{T}$$

- At the beginning of each period t the agent observes the realization (x_t, ϵ_t) chooses d_t to sequentially maximize:

$$E\left\{ \sum\nolimits_{\tau=t}^{T} \sum\nolimits_{j=1}^{J} \beta^{\tau-1} d_{j\tau}\left[u_{j\tau}(x_\tau) + \epsilon_{j\tau}\right] | x_t, \epsilon_t \right\} \quad (1)$$

where the expectation is taken over future realized values $x_{t+1}, \ldots, x_T$ and $\epsilon_{t+1}, \ldots, \epsilon_T$ conditional on (x_t, ϵ_t).

A Class of Dynamic Discrete Choice Markov Models

Optimization

- Denote the optimal decision rule at t as $d_t^o(x_t, \epsilon_t)$, with j^{th} element $d_{jt}^o(x_t, \epsilon_t)$ and define:

$$V_t(x_t) \equiv E\left\{\sum_{\tau=t}^{T}\sum_{j=1}^{J}\beta^{\tau-t-1}d_{j\tau}^o\left(x_\tau, \epsilon_\tau\right)\left(u_{j\tau}(x_\tau) + \epsilon_{j\tau}\right)\right\}$$

- The conditional value function, $v_{jt}(x_t)$, is defined as:

$$v_{jt}(x_t) = u_{jt}(x_t) + \beta\sum_{x=1}^{X}V_{t+1}(x)f_{jt}\left(x|x_t\right)$$

- Integrating $d_{jt}^o(x_t, \epsilon)$ over $\epsilon \equiv (\epsilon_1, \ldots, \epsilon_J)$:

$$p_{jt}(x_t) \equiv E\left[d_{jt}^o\left(x_t, \epsilon\right)|x_t\right] = \int d_{jt}^o\left(x_t, \epsilon\right)g_t\left(\epsilon\,|x_t\right)d\epsilon$$

- The starting point for our analysis is to define differences in the conditional valuation functions as:

$$\Delta v_{jkt}(x) \equiv v_{jt}(x) - v_{kt}(x)$$

- Although there are $J(J-1)$ differences all but $(J-1)$ are linear combinations of the $(J-1)$ basis functions.
- For example setting the basis functions as:

$$\Delta v_{jt}(x) \equiv v_{jt}(x) - v_{Jt}(x)$$

  then clearly:

$$\Delta v_{jkt}(x) = \Delta v_{jt}(x) - \Delta v_{kt}(x)$$

- Without loss of generality we focus on this particular basis function.

Inversion

Each CCP is a mapping of differences in the conditional valuation functions

- Using the definition of $\Delta v_{jt}(x)$:

$$
\begin{aligned}
p_{jt}(x) &\equiv \int d^o_{jt}(x,\epsilon)\, g_t(\epsilon\,|x)\, d\epsilon \\
&= \int I\left\{\epsilon_k \leq \epsilon_j + \Delta v_{jt}(x) - \Delta v_{kt}(x) \forall\ k \neq j\right\} g_t(\epsilon\,|x)\, d\epsilon \\
&= \int_{-\infty}^{\epsilon_j+\Delta v_{jt}(x)-\Delta v_{1t}(x)} \cdots \int_{-\infty}^{\epsilon_j+\Delta v_{jt}(x)-\Delta v_{J-1,t}(x)} \int_{-\infty}^{\epsilon_j+\Delta v_{jt}(x)} g_t(\epsilon\,|x)\, d\epsilon
\end{aligned}
$$

- Noting $g_t(\epsilon\,|x) \equiv \partial^J G_t(\epsilon\,|x)/\partial\epsilon_1,\ldots,\partial\epsilon_J$, integrate over $(\epsilon_1,\ldots,,\epsilon_{j-1},\epsilon_{j+1}\ldots,\epsilon_J)$.
- Denoting $G_{jt}(\epsilon\,|x) \equiv \partial G_t(\epsilon\,|x)/\partial\epsilon_j$, yields:

$$
p_{jt}(x) = \int_{-\infty}^{\infty} G_{jt}\left(\begin{array}{c} \epsilon_j + \Delta v_{jt}(x) - \Delta v_{1t}(x), \ldots \\ \ldots, \epsilon_j, \ldots, \epsilon_j + \Delta v_{jt}(x) \end{array} |x \right) d\epsilon_j
$$

Inversion

There are as many CCPs as there are conditional valuation functions

- For any vector $J-1$ dimensional vector $\delta \equiv (\delta_1, \ldots, \delta_{J-1})$ define:

$$Q_{jt}\left(\delta, x\right) \equiv \int_{-\infty}^{\infty} G_{jt}\left(\epsilon_j + \delta_j - \delta_1, \ldots, \epsilon_j, \ldots, \epsilon_j + \delta_j \left| x \right.\right) d\epsilon_j$$

- We interpret $Q_{jt}\left(\delta, x\right)$ as the probability taking action j in a static random utility model (RUM) where the payoffs are $\delta_j + \epsilon_j$ and the probability distribution of disturbances is given by $G_t\left(\epsilon \left| x \right.\right)$.
- It follows from the definition of $Q_{jt}\left(\delta, x\right)$ that:

$$0 \leq Q_{jt}\left(\delta, x\right) \leq 1 \text{ for all } \left(j, t, \delta, x\right) \text{ and } \sum_{j=1}^{J-1} Q_{jt}\left(\delta, x\right) \leq 1$$

- In particular the previous slide implies that for any given $\left(j, t, x\right)$:

$$p_{jt}(x) = \int_{-\infty}^{\infty} G_{jt}\left(\begin{array}{c} \epsilon_j + \Delta v_{jt}(x) - \Delta v_{1t}(x), \\ \ldots, \epsilon_j, \ldots, \epsilon_j + \Delta v_{jt}(x) \end{array} \left| x \right.\right) d\epsilon_j \equiv Q_{jt}\left(\Delta v_t(x), x\right)$$

Theorem (Inversion)

For each (t, δ, x) *define:*

$$Q_t(\delta, x) \equiv (Q_{1t}(\delta, x), \ldots Q_{J-1,t}(\delta, x))'$$

Then the vector function $Q_t(\delta, x)$ *is invertible in* δ *for each* (t, x).

- Note that $p_{Jt}(x) = Q_{Jt}(\Delta v_t, x)$ is a linear combination of the other equations in the system because $\sum_{k=1}^{J} p_k = 1$.
- Let $p \equiv (p_1, \ldots, p_{J-1})$ where $0 \leq p_j \leq 1$ for all $j \in \{1, \ldots, J-1\}$ and $\sum_{j=1}^{J-1} p_j \leq 1$. Denote the inverse of $Q_{jt}(\Delta v_t, x)$ by $Q_{jt}^{-1}(p, x)$.
- The inversion theorem implies:

$$\begin{bmatrix} \Delta v_{1t}(x) \\ \vdots \\ \Delta v_{J-1,t}(x) \end{bmatrix} = \begin{bmatrix} Q_{1t}^{-1}[p_t(x), x] \\ \vdots \\ Q_{J-1,t}^{-1}[p_t(x), x] \end{bmatrix}$$

Inversion

Using the inversion theorem

- In finessing optimization and integration by exploiting conditional independence, how far can the three applications described in the previous two lectures be extended?
- We use the Inversion Theorem to:
  1. provide empirically tractable representations of the conditional value functions.
  2. analyze identification in dynamic discrete choice models.
  3. provide convenient parametric forms for the density of ϵ_t that generalize the Type 1 Extreme Value distribution.
  4. generalize the renewal and terminal state properties exploited in the first two examples, by obtaining restrictions on the state variable transitions used to implement CCP estimators.
  5. introduce new methods for incorporating unobserved state variables.

Representation and Identification

Identifying the policy function

- From the definition of the optimal decision rule, and appealing to the inversion theorem, it is easy to show:

$$d_{jt}^{o}\left(x_t,\epsilon_t\right)=\prod\nolimits_{k=1}^{J}1\left\{\epsilon_{kt}-\epsilon_{jt}\leq Q_{jt}^{-1}\left[p_t(x),x\right]-Q_{kt}^{-1}\left[p_t(x),x\right]\right\}$$

- If $G_t\left(\epsilon\,|x\right)$ is known and the data generating process (DGP) is $\left(x_t,d_t\right)$, then $p_t(x)$ and hence $d_t^{o}\left(x_t,\epsilon_t\right)$ are identified.
- For example if ϵ_{jt} is distributed IID T1EV, then:

$$d_{jt}^{o}\left(x_t,\epsilon_t\right)=\prod\nolimits_{k=1}^{J}1\left\{\epsilon_{kt}-\epsilon_{jt}\leq\ln\left[p_{jt}(x)\right]-\ln\left[p_{kt}(x)\right]\right\}$$

- Define the conditional value function correction as:

$$\psi_{jt}(x) \equiv V_t(x) - v_{jt}(x)$$

- Suppose the agent committed to taking action j before seeing ϵ_t:
  - The expected lifetime utility would be $v_{jt}(x_t) + E_t\left[\epsilon_{jt} \,|x_t\right]$
  - It entails a loss of $\psi_{jt}(x) - E_t\left[\epsilon_{jt} \,|x_t\right]$, simplifying to $\psi_{jt}(x)$ if $E_t\left[\epsilon_{jt} \,|x_t\right] = 0$.
- One can show $\psi_{jt}(x)$ is also identified (Arcidiacono and Miller, 2011).
- For example if $G(\epsilon)$ is distributed nested logit:

$$G(\epsilon) \equiv \exp\left[-\left(\sum_{j=1}^{J} \exp\left[-\epsilon_j/\sigma\right]\right)^{\sigma}\right]$$

where $\sigma \in [0, 1]$ is the correlation within the nest, then:

$$\psi_j(p) = \gamma - \sigma \ln(p_j) - (1-\sigma)\ln\left(\sum_{k=1}^{J} p_k\right)$$

- From its definition:

$$v_{jt}(x_t) = u_{jt}(x_t) + \beta \sum_{x=1}^{X} V_{t+1}(x) f_{jt}(x_{t+1}|x_t)$$

- Substituting for $V_{t+1}(x_{t+1})$ using conditional value function correction we obtain for any k:

$$v_{jt}(x_t) = u_{jt}(x_t) + \beta \sum_{x=1}^{X} \left[v_{k,t+1}(x) + \psi_{k,t+1}(x) \right] f_{jt}(x|x_t)$$

- We could repeat this procedure ad infinitum, substituting in for $v_{k,t+1}(x)$ by using the definition for $\psi_{kt}(x)$.

- For example assume $u_{Jt}(x) \equiv 0$, and that β and G are known.
- For all (t, x) and define:

$$\kappa_\tau(x_{\tau+1}|x_t, j) \equiv \begin{cases} f_{jt}(x_{t+1}|x_t) \text{ if } \tau = t \\ \sum_{x_\tau=1}^{X} f_{1\tau}(x_{\tau+1}|x_\tau)\kappa_{\tau-1}(x_\tau|x_t, j) \text{ if } \tau > t \end{cases}$$

- Unravelling $V_{t+1}(x)$ and $v_{1,t+1}(x)$ yields:

$$v_{jt}(x_t) = u_{jt}(x_t) + \sum_{\tau=t+1}^{T} \sum_{x=1}^{X} \beta^{\tau-t} \psi_J[p_\tau(x)]\kappa_{\tau-1}(x|x_t, j)$$

- Noting $\Delta v_{jt}(x) = Q_{jt}^{-1}[p_t(x), x]$ proves $u_{jt}(x_t)$ is identified because:

$$u_{jt}(x_t) = \sum_{\tau=t+1}^{T} \sum_{x=1}^{X} \beta^{\tau-t} \psi_J[p_\tau(x)] \begin{bmatrix} \kappa_{\tau-1}(x|x_t, J) \\ -\kappa_{\tau-1}(x|x_t, j) \end{bmatrix} - Q_{jt}^{-1}[p_t(x), x]$$

- Two of the earliest papers estimating dynamic discrete choice models, Miller (1984) and Rust (1987), exploit the *renewal* property:
  - an action in the choice set returning the state to its initial conditions.
- Mr. Zurcher maximizes the expected discounted sum of payoffs:

$$E\left\{\sum_{t=1}^{\infty}\beta^{t-1}\left[d_{t2}(\theta_1 x_t+\theta_2 s+\epsilon_{t2})+d_{t1}\epsilon_{t1}\right]\right\}$$

  - where s is an invariant bus characteristic (model)
  - and x_{t+1} is mileage on the engine, evolving deterministically as:

$$x_{t+1}=\begin{cases}1 & \text{if } j=1\\ x_t+1 & \text{if } j=2\end{cases}$$

- Thus setting $j=1$ is the renewal action.

A Renewal Problem

The value function and optimal decision rule

- This is a *stationary infinite horizon problem*, so age and time have no role.
- Denote by $d_1^o(x, s, \epsilon_t) = 1 - d_2^o(x, s, \epsilon_t)$ whether to optimally renew (replace the engine).
- Also denote the ex-ante value, or *social surplus*, function by:

$$V(1, s) = E\left\{\sum_{t=1}^{\infty} \beta^{t-1}\left[d_{t2}(\theta_1 x_t + \theta_2 s + \epsilon_{t2}) + d_{t1}\epsilon_{t1}\right]\right\}$$

- Then, appealing to Bellman's principle, the conditional value function for each choice is:

$$v_j(x, s) = \begin{cases} \beta V(1, s) & \text{if } j = 1 \\ \theta_1 x + \theta_2 s + \beta V(x + 1, s) & \text{if } j = 2 \end{cases}$$

and optimizing behavior implies:

$$d_1^o(x, s, \epsilon_t) = \mathbf{1}\{\epsilon_{t2} - \epsilon_{t1} \leq v_1(x, s) - v_2(x, s)\} = 1 - d_2^o(x, s, \epsilon_t)$$

- We suppose the data comprises a cross section of N observations of buses $n \in \{1, \ldots, N\}$ reporting their:
  - fixed characteristics s_n,
  - engine miles x_n,
  - and maintenance decision (d_{n1}, d_{n2}).
- Let $p_1(x, s)$ denote the conditional choice probability (CCP) of replacing the engine given x and s.
- Stationarity and T1EV imply that for all t :

$$\begin{aligned} p_1(x, s) &\equiv \int_{\epsilon_t} d_1^o(x, s, \epsilon_t) g(\epsilon_t) d\epsilon_t \\ &= \int_{\epsilon_t} \mathbf{1}\{\epsilon_{t2} - \epsilon_{t1} \leq v_1(x, s) - v_2(x, s)\} g(\epsilon_t | x_t) d\epsilon_t \\ &= \{1 + \exp[v_2(x, s) - v_1(x, s)]\}^{-1} \end{aligned}$$

- An ML estimator could be formed off this equation.

A Renewal Problem

Exploiting the renewal property

- Rust showed that if ϵ_{jt} is T1EV, then for all $(x, s,)$:

$$V(x,s) = v_j(x,s) - \beta \log [p_j(x,s)] + 0.57\ldots$$

- Therefore the conditional value function of not replacing is:

$$\begin{aligned} v_2(x,s) &= \theta_1 x + \theta_2 s + \beta V(x, s+1) \\ &= \theta_1 x + \theta_2 s + \beta \{ v_1 (x+1, s) - p_1(x+1, s) + 0.57\ldots \} \end{aligned}$$

- Similarly:

$$v_1(x,s) = \beta V(1,s) = \beta \{ v_1(1,s) - \ln [p_1(1,s)] + 0.57 \} \ldots$$

- Because bus engine miles is the only factor affecting bus value given s:

$$v_1(x+1, s) = v_1(1, s)$$

A Renewal Problem

Using CCPs to represent differences in continuation values

- Hence:

$$v_2(x, s) - v_1(x, s) = \theta_1 x + \theta_2 s + \beta \ln [p_1(1, s)] - \beta \ln [p_1(x + 1, s)]$$

- Therefore:

$$\begin{aligned} p_1(x, s) &= \frac{1}{1 + \exp [v_2(x, s) - v_1(x, s)]} \\ &= \frac{1}{1 + \exp \left\{ \theta_1 x + \theta_2 s + \beta \ln \left[\frac{p_1(1,s)}{p_1(x+1,s)} \right] \right\}} \end{aligned}$$

- Intuitively the CCP for current replacement is the CCP for a static model with an offset term.
- The offset term accounts for differences in continuation values using future CCPs that characterize optimal future replacements.

- Consider the following CCP estimator:
  1. Form a first stage estimator for $p_1(x,s)$ from the relative frequencies:
  $$\hat{p}_1(x,s) \equiv \frac{\sum_{n=1}^{N} d_{n1} I(x_n = x) I(s_n = s)}{\sum_{n=1}^{N} I(x_n = x) I(s_n = s)}$$
  2. Substitute $\hat{p}_1(x,s)$ into the likelihood as incidental parameters to estimate $(\theta_1, \theta_2, \beta)$ with a logit:
  $$\frac{d_{n1} + d_{n2} \exp(\theta_1 x_n + \theta_2 s_n + \beta \ln\left[\frac{\hat{p}_1(1,s_n)}{\hat{p}_1(x_n+1,s_n)}\right]}{1 + \exp(\theta_1 x_n + \theta_2 s_n + \beta \ln\left[\frac{\hat{p}_1(1,s_n)}{\hat{p}_1(x_n+1,s_n)}\right]}$$
  3. Correct the standard errors for $(\theta_1, \theta_2, \beta)$ induced by the first stage estimates of $p_1(x,s)$.
- Note that in the second stage $\ln\left[\frac{\hat{p}_1(1,s_n)}{\hat{p}_1(x_n+1,s_n)}\right]$ enters the logit as an individual specific component of the data, the β coefficient entering in the same way as θ_1 and θ_2.

Monte Carlo Study (Arcidiacono and Miller, 2011)

Modifying the bus engine problem

- Suppose bus type $s \in \{0, 1\}$ is equally weighted.
- Two state variables affect wear and tear on the engine:
  1. total accumulated mileage:

$$x_{1,t+1} = \begin{cases} \Delta_t \text{ if } d_{1t} = 1 \\ x_{1t} + \Delta_t \text{ if } d_{2t} = 1 \end{cases}$$

  2. a permanent route characteristic for the bus, x_2, that systematically affects miles added each period.
- More specifically we assume:
  - $\Delta_t \in \{0, 0.125, \ldots, 24.875, 25\}$ is drawn from a discretized truncated exponential distribution, with:

$$f(\Delta_t | x_2) = \exp\left[-x_2(\Delta_t - 25)\right] - \exp\left[-x_2(\Delta_t - 24.875)\right]$$

  - x_2 is a multiple 0.01 drawn from a discrete equi-probability distribution between 0.25 and 1.25.

Monte Carlo Study

Including the age of the bus in panel estimation

- Let θ_{0t} denote other bus maintenance costs tied to its vintage.
- This modification renders the optimization problem nonstationary.
- The payoff difference from retaining versus replacing the engine is:

$$u_{t2}(x_{t1}, s) - u_{t1}(x_{t1}, s) \equiv \theta_{0t} + \theta_1 \min\{x_{t1}, 25\} + \theta_2 s$$

- Denoting $x_t \equiv (x_{1t}, x_2)$, this implies:

$$\begin{aligned} v_{t2}(x_t, s) - v_{t1}(x_t, s) &= \theta_{0t} + \theta_1 \min\{x_{t1}, 25\} + \theta_2 s \\ &\quad + \beta \sum_{\Delta_t \in \Lambda} \left\{ \ln\left[\frac{p_{1t}(\Delta_t, s)}{p_{1t}(x_{1t} + \Delta_t, s)} \right] \right\} f(\Delta_t | x_2) \end{aligned}$$

- In the first three columns of the next table each sample simulation has 1000 buses observed for 20 periods.
- In the fourth column 2000 buses are observed for 10 periods.
- The mean and standard deviations are compiled from 50 simulations.

Monte Carlo Study

Extract from Table 1 of Arcidiacono and Miller (2011)

Monte Carlo for Optimal Stopping Problem+

	DGP (1)	FIML (2)	CCP (3)	Time effects CCP (4)
θ_0 (Intercept)	2	2.0100	1.9911	
		(0.0405)	(0.0399)	
θ_1 (Mileage)	-0.15	-0.1488	-0.1441	-0.1440
		(0.0074)	(0.0098)	(0.0121)
θ_2 (Type)	1	0.9945	0.9726	0.9683
		(0.0611)	(0.0668)	(0.0636)
β (Discount Factor)	0.9	0.9102	0.9099	0.9172
		(0.0411)	(0.0554)	(0.0639)
Time (Minutes)		130.29	0.078	0.079
		(19.73)	(0.0041)	(0.0047)

+ Mean and standard deviations for fifty simulations. For columns (2) and (3), the observed data consist of 1000 buses for 20 periods. For column (4), the intercept (θ_0) is allowed to vary over time and the data consist of 2000 buses for 10 periods.

- The basic difference between estimating a dynamic game and an individual optimization problem using a CCP estimator revolves around how much the payoffs of each player are affected by state variables partially determined by other players through their conditional choice probabilities.
- Note that:
  1. there might be multiple equilibria, but we assume:
     - either every firm plays in the same market
     - or every market plays the same equilibrium.
  2. in contrast to ML we do not solve for the equilibrium.
  3. estimation is based on conditions that are satisfied by every equilibrium.
  4. the estimation approach is identical to the approach we described in the individual optimization problem.

- To convey some flavor of how CCP estimation extends to games consider the following.
- Suppose there is a finite maximum number of firms in a market at any one time denoted by I.
- If a firm exits, the next period an opening occurs to a potential entrant, who may decide to exercise this one time option, or stay out.
- At the beginning of each period every incumbent firm has the option of quitting the market or staying one more period.
- Let $d_t^{(i)} \equiv \left(d_{t1}^{(i)}, d_{t2}^{(i)} \right)$, where $d_{t1}^{(i)} = 1$ means i exits or stays out of the market in period t, and $d_{t2}^{(i)} = 1$ means i enters or does not exit.
- If $d_{t2}^{(i)} = 1$ and $d_{t-1,1}^{(i)} = 1$ then the firm in spot i at time t is an entrant, and if $d_{t-1,2}^{(i)} = 1$ the spot i at time t is an incumbent.

- In this application there are three components to the state variables and $x_t = (x_1, x_{2t}, s_t)$.
- The first is a permanent market characteristic, denoted by x_1, and is common across firms in the market. Each market faces an equal probability of drawing any of the possible values of x_1 where $x_1 \in \{1, 2, \ldots, 10\}$.
- The second, x_{2t}, is whether or not each firm is an incumbent, $x_{2t} \equiv \{d_{t-1,2}^{(1)}, \ldots, d_{t-1,2}^{(I)}\}$. Entrants pay a start up cost, making it more likely that stayers choose to fill a slot than an entrant.
- A demand shock $s_t \in \{1, \ldots, 5\}$ follows a first order Markov chain.
- In particular, the probability that $s_{t+1} = s_t$ is fixed at $\pi \in (0, 1)$, and probability of any other state occurring is equally likely:

$$\Pr\left\{s_{t+1} \left| s_t \right.\right\} = \left\{ \begin{array}{c} \pi \text{ if } s_{t+1} = s_t \\ (1-\pi)/4 \text{ if } s_{t+1} \neq s_t \end{array} \right.$$

- Each active firm produces one unit so revenue, denoted by y_t, is just price.
- Price is determined by:
  1. the supply of active firms in the market, $\sum_{i=1}^{I} d_{t2}^{(i)}$
  2. a permanent market characteristic, x_1
  3. the Markov demand shock s_t
  4. another temporary shock, denoted by η_t, distributed *iID* standard normal distribution, revealed to each market after the entry and exit decisions are made.
- The price equation is:

$$y_t = \alpha_0 + \alpha_1 x_1 + \alpha_2 s_t + \alpha_3 \sum_{i=1}^{I} d_{t2}^{(i)} + \eta_t$$

- We assume costs comprise a choice specific disturbance $\epsilon_{tj}^{(i)}$ that is privately observed, plus a linear function of $\left(x_t^{(i)}, s_t^{(i)}, d_t^{(-i)}\right)$.
- Net current profits for exiting incumbent firms, and potential entrants who do not enter, are $\epsilon_{1t}^{(i)}$. Thus $U_1^{(i)}\left(x_t^{(i)}, s_t^{(i)}, d_t^{(-i)}\right) \equiv 0$.
- Current profits from being active are the sum of $\left(\epsilon_{2t}^{(i)} + \eta_t\right)$ and:

$$U_2^{(i)}\left(x_t^{(i)}, s_t^{(i)}, d_t^{(-i)}\right) \equiv \theta_0 + \theta_1 x_1 + \theta_2 s_t + \theta_3 \sum_{\substack{i'=1 \\ i' \neq i}}^{I} d_{2t}^{(i')} + \theta_4 d_{1,t-1}^{(i)}$$

  where θ_4 is the startup cost that only entrants pay.
- In equilibrium $E\left(\eta_t\right) = 0$ so:

$$u_j^{(i)}\left(x_t, s_t\right) = \theta_0 + \theta_1 x_1 + \theta_2 s_t + \theta_3 \sum_{\substack{i'=1 \\ i' \neq i}}^{I} p_2^{(i')}\left(x_t, s_t\right) + \theta_4 d_{1,t-1}^{(i)}$$

- We assume the firm's private information, $\epsilon_{jt}^{(i)}$, is distributed T1EV.
- In this model exiting is a *terminal choice*, with a payoff normalized to zero, given T1EV, the conditional value function for being active is:

$$\begin{aligned} v_2^{(i)}(x_t, s_t) &= u_2^{(i)}(x_t, s_t) \\ &\quad -\beta \sum_{x \in X} \sum_{s \in S} \left(\ln \left[p_1^{(i)}(x, s) \right] \right) f_2^{(i)}(x, s | x_t, s_t) \end{aligned}$$

- The future value term is then expressed as a function solely of the one-period-ahead probabilities of exiting and the transition probabilities of the state variables.

- The number of firms in each market is set to six and we simulated data for 3,000 markets.
- The discount factor is set to $\beta = 0.9$.
- Starting at an initial date with six potential entrants in the market, we solved the model, ran the simulations forward for twenty periods, and used the last ten periods to estimate the model.
- The key difference between this Monte Carlo and the renewal Monte Carlo is that the conditional choice probabilities have an additional effect on both current utility and the transitions on the state variables due to the effect of the choices of the firm's competitors on profits.

Entry Exit Game

Extract from Table 2 of Arcidiacono and Miller (2011)

	DGP (1)	s_t Observed (2)
Profit parameters		
θ_0 (intercept)	0	0.0207 (0.0779)
θ_1 (obs. state)	0.05	−0.0505 (0.0028)
θ_2 (unobs. state)	0.25	0.2529 (0.0080)
θ_3 (no. of competitors)	−0.2	−0.2061 (0.0207)
θ_4 (entry cost)	−1.5	−1.4992 (0.0131)
Price parameters		
α_0 (intercept)	7	6.9973 (0.0296)
α_1 (obs. state)	−0.1	−0.0998 (0.0023)
α_2 (unobs. state)	0.3	0.2996 (0.0045)
α_3 (no. of competitors)	−0.4	−0.3995 (0.0061)
π (persistence of unobs. state)	0.7	
Time (minutes)		0.1354 (0.0047)

- This coming academic year I will give:
  1. 14 lectures in the fall on concepts used in structural econometrics. Topics include: summarizing the data, elementary probability theory, convergence, estimators, and hypothesis testing. To motivate these concepts, we begin the series with several applications.
  2. 14 lectures in the spring on (i) dynamic discrete choice estimation that expands on what I have presented today with applications in technological change and industrial organization (7 lectures), and (ii) on estimating structural models of market microstructure, including auctions, limit order markets and optimal contracting (7 lectures).
- You are welcome to sign up for these free lectures by emailing me at:

ramiller@cmu.edu.

- The lecture series webpage, including the schedule, is posted at:

http://comlabgames.com/structuraleconometrics/.

Additional Instruction

Remote courses

- In conjunction with these two sets of lectures, Helsinki GSE is offering two courses, Structural Econometrics 1 and 2, globally to Ph.D. students and postdoctoral researchers, as part of their public mission to making higher education more accessible:
    1. Both courses are based on and coordinated with the material presented in the public lectures; they consist of tutorials, several assignments and possibly a take-home final examination.
    2. They can be taken as stand alone courses, or if your home institution permits, may be used as certification for an elective taken (remotely) at Helsinki GSE.
    3. See https://www.helsinkigse.fi/research/ and https://www.helsinkigse.fi/courses/.
    4. The courses are offered at no charge, but currently the class size is limited to 22 students, so there may be nonprice rationing for the 22 seats currently available.
    5. The instructor and contact person for these courses is Ciprian Domnisoru https://www.cipriandomnisoru.net/