Dynamic Structural Models

Bus Engine CCP

Aaron Barkley

Aug 2025

Static: McFadden's Random Utility Model

- Binary choice model of random utility:

$$U_j(x, \epsilon_j) = u_j(x) + \epsilon_j, \quad j \in \{0, 1\}$$

  where
  - x is an observed state variable,
  - u_j is the deterministic component of utility,
  - ϵ_j is an iid random utility shock for each choice $j = 0, 1$. Observed to the agent but not by the econometrician.

- Agent's goal: choose the option with the highest utility, i.e., take action 1 iff $U_1(x, \epsilon_1) > U_0(x, \epsilon_0)$.

- Econometrician's goal: estimate u_j given data on choices d and states: $\{d_n, x_n\}_{n=1}^N$.

Random utility model estimation

- Normalize one of the choices: set $u_0(x) = 0$ for all x.
- Parameterize utility for the other choice: $u_1(x) = \theta_0 + \theta_1 x$.
- Assume that $\epsilon_j \overset{iid}{\sim}$ Type I EV for all j.
- Choice one is made $(d = 1)$ in state x iff

$$
\begin{aligned}
\theta_0 + \theta_1 x + \epsilon_1 &> \epsilon_0 \\
\Rightarrow \quad \mathrm{Pr}(d = 1|x) &= \mathrm{Pr}(\epsilon_0 - \epsilon_1 < \theta_0 + \theta_1 x) \\
&= \frac{\exp(\theta_0 + \theta_1 x)}{1 + \exp(\theta_0 + \theta_1 x)}
\end{aligned}
$$

  because $\epsilon_1 - \epsilon_0 \sim$Logistic(0,1) when both $\epsilon \sim$Type I EV.

- Estimate θ_0, θ_1 with a logistic regression.

From static to dynamic

- Now we move to a dynamic environment: discount factor β and state transition probabilities F_j. Consumer chooses action j to maximize expected discounted future utility.

- Binary choice. Take action $j = 1$ iff

$$v_1(x_t) + \epsilon_{1t} > v_0(x_t) + \epsilon_{0t}$$

- Maintain the assumption that $\epsilon_{jt} \sim$ Type I EV for all j, t. Probability that decision in time t is choice one, or $\text{Pr}(d_t = 1)$, is

$$\text{Pr}(d_t = 1) = \frac{\exp(v_1(x_t) - v_0(x_t))}{1 + \exp(v_1(x_t) - v_0(x_t))}$$

Introduction

- Problem: we cannot simply parameterize v_j and estimate directly with a logit. That would give us parameters that are not related to the theory.
- Recall that

$$v_j(x_t) = u_j(x_t) + \beta \sum_{x=1}^{\overline{x}} V(x_{t+1}) f_j(x_{t+1} \, | x_t)$$

and

$$V_t(x_t) \equiv E\left\{ \sum_{\tau=t}^{\overline{T}} \sum_{j=0}^{1} \beta^{\tau-t-1} d^o_{jt}(x_\tau \epsilon_\tau)(u_{j\tau}(x_\tau) + \epsilon^*_{j\tau}) \right\}$$

where d^0_{jt} represents that choice j was optimal in time t, and ϵ^*_{jt} is the value of ϵ conditional on j being optimal.

The CCP solution

- If we knew ϵ^*_{jt}, then we could express $V_t(x_t)$ in terms of utility parameters and things that we already know or can estimate.
- The insight of Hotz and Miller (1993) is that we can *estimate* ϵ^*_{jt} *using the conditional choice probabilities*!
- Specifically, if $\epsilon_{jt} \sim$ Type I EV, then

$$\epsilon^*_{jt} = \gamma - \ln(p_j(x_t)), \quad V(x) = v_j(x) + \gamma - \ln(p_j(x_t))$$

- Now the conditional value function can be expressed as

$$v_j(x_t) = u_j(x_t) + \beta \sum_{x_{t+1} \in \mathcal{X}} f_j(x_{t+1}|x_t) \left[v_0(x_{t+1}) + \gamma - \ln(p_0(x_{t+1}))\right]$$

The CCP solution

- We continue to replace v_0 with a u_0 term and a $V(x_{t+1})$ term.
- Every time we see $V(x_\tau)$, we replace it with $v_0(x_\tau) + \gamma - \ln(p_0(x_\tau))$. Then the right-hand side will only have F_j, F_0, u_0, and $\gamma - \ln(p_0(x_\tau))$ terms.
- With some matrix algebra and setting $\overline{T} = \infty$, the difference between the conditional value functions is:

$$\begin{aligned}
\mathbf{v}_1 - \mathbf{v}_0 &= \mathbf{u}_1 - \mathbf{u}_0 + \beta\Big(F_1 - F_0\Big)\Big[I - \beta F_0\Big]^{-1}\Big(\Psi_0 + \mathbf{u}_0\Big) \\
&= X\theta + \underbrace{\beta\Big(F_1 - F_0\Big)\Big[I - \beta F_0\Big]^{-1}\Big(\Psi_0\Big)}_{\equiv \Delta\mathbf{V}}
\end{aligned}$$

after using $u_0(x) = 0$ for all x and linear utility for u_1.

Estimation steps with linear utility:

1. Estimate the CCPs $p_j(x_{jt})$ and state transitions f_1, f_0.
2. Calculate the differenced "Big V" term:

$$\Delta\mathbf{V} = \beta\Big(F_1 - F_0\Big)\Big[I - \beta F_0\Big]^{-1}\Big(\gamma - \ln(\hat{\mathbf{p}}_0)\Big)$$

3. Specify logistic log-likelihood function Q:

$$\begin{aligned} \Delta\mathbf{v} &\equiv \mathbf{v}_1 - \mathbf{v}_0 = X\theta + \Delta\mathbf{V} \\ Q(\theta) &= \sum_{n=1}^{N}\sum_{t=1}^{T}\mathbf{1}\{d_{nt} = 1\}\Delta v(x_{nt}) - \ln(1 + \exp(\Delta v(x_{nt}))) \end{aligned}$$

4. Maximize $Q(\theta)$ over θ.