Dynamic Structural Models

2. Bus Engine CCP

Aaron Barkley
Department of Economics

July 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

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

  because $\epsilon_1 - \epsilon_0$ ~Logistic(0,1) when both ϵ ~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 $\Pr(d_t = 1)$, is

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

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

$$v_j(x_t) = u_j(x_t) + \beta \sum_{x=1}^{\bar{x}} V(x_{t+1}) f_1(x_{t+1} \,|x_t)$$

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

$$\epsilon^*_{jt} = \gamma - \ln(p_j(x_t))$$

- Now the value function can be expressed as

$$V(x_t) = \sum_{\tau=t}^{\overline{T}} \sum_{x_{t+1} \in \mathcal{X}} \sum_{j=0}^{1} \beta^{\tau-t-1} f_j(x_{\tau+1}|x_\tau) \left[p_j(x_\tau)(u_j(x_\tau) + \gamma - \ln(p_j(x_\tau)))\right]$$

The CCP solution

- With some matrix algebra and setting $\overline{T} = \infty$ and assuming $U_1(X) = X\theta$ and and $U_0(X) = 0$, the ex-ante value function is:

$$\mathbf{V} = \underbrace{\left(I - \beta\sum_j (\mathbf{p}_j \lambda) * F_j\right)^{-1}}_{\equiv A} \left(\underbrace{\mathbf{p}_1 * (\gamma - \ln(\mathbf{p}_1)) + \mathbf{p}_0 * (\gamma - \ln(\mathbf{p}_0))}_{\equiv B} + \underbrace{\mathbf{p}_1 * X}_{\equiv C}\theta\right)$$

  where $*$ is element-wise multiplication and λ is a vector of ones.

- Recall that we want $v_1(x) - v_0(x)$ to compute the logit likelihood:

$$\mathbf{v}_1 - \mathbf{v}_0 = \mathbf{u}_1 - \mathbf{u}_0 + \beta(F_1 - F_0)\mathbf{V}$$

  Everything on the right-hand side is a known function of θ and the data.

Estimation steps with linear utility:

1. Estimate the CCPs $p_j(x_{jt})$ and state transitions f_1, f_0.
2. Calculate three matrices/vectors:
   - $A = \text{inv}(I - \beta \sum_j (\mathbf{p}_j \lambda) * F_j)$
   - $B = \mathbf{p}_1 * (\gamma - \ln(\mathbf{p}_1)) + \mathbf{p}_0 * (\gamma - \ln(\mathbf{p}_0))$
   - $C = \mathbf{p}_1 * X$
3. Specify logistic likelihood function Q that takes as inputs θ, A, B, C:

$$\begin{aligned} \Delta \mathbf{v} &\equiv \mathbf{v}_1 - \mathbf{v}_0 = X\theta + \beta(F_1 - F_0) \times [A(B + C\theta)] \\ Q(\theta) &= \sum_{t=1}^{T} \left(\frac{\exp(\Delta v(x_t))}{1 + \exp(\Delta v(x_t))} \right)^{d_t} \times \left(\frac{1}{1 + \exp(\Delta v(x_t))} \right)^{1-d_t} \end{aligned}$$

4. Maximize $\ln(Q)$ over θ.