Dynamic Selection Problem

Link choices to outcomes:

- Labor force participation $\rightarrow$ earnings
- Market entry $\rightarrow$ profits

Dynamic Selection Problem

Link choices to outcomes:

- Labor force participation $\rightarrow$ earnings
- Market entry $\rightarrow$ profits

Selection bias: firms choose to enter a market based on unobservable potential profits

We can use finite mixture distributions to control for selection

Dynamic Selection Problem

Link choices to outcomes:

- Labor force participation $\rightarrow$ earnings
- Market entry $\rightarrow$ profits

Selection bias: firms choose to enter a market based on unobservable potential profits

We can use finite mixture distributions to control for selection

Assumption: Selection bias eliminated once we control for unobserved type

Let Y_{2t} = choice, Y_{1t} = outcome, s is unobserved type. Joint likelihood given by

Let Y_{2t} = choice, Y_{1t} = outcome, s is unobserved type. Joint likelihood given by

$$\mathcal{L}(Y_{1t}, Y_{2t}|X_{1t}, X_{2t}, \alpha_1, \alpha_2, s) = \mathcal{L}(Y_{1t}|Y_{2t}, X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s)$$

Let Y_{2t} = choice, Y_{1t} = outcome, s is unobserved type. Joint likelihood given by

$$\begin{aligned}\mathcal{L}(Y_{1t}, Y_{2t}|X_{1t}, X_{2t}, \alpha_1, \alpha_2, s) &= \mathcal{L}(Y_{1t}|Y_{2t}, X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s) \\ &= \mathcal{L}(Y_{1t}|X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s)\end{aligned}$$

Let Y_{2t} = choice, Y_{1t} = outcome, s is unobserved type. Joint likelihood given by

$$
\begin{aligned}
\mathcal{L}(Y_{1t}, Y_{2t}|X_{1t}, X_{2t}, \alpha_1, \alpha_2, s) &= \mathcal{L}(Y_{1t}|Y_{2t}, X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s) \\
&= \mathcal{L}(Y_{1t}|X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s)
\end{aligned}
$$

Likelihood separable conditional on type s, i.e. $Y_{1t} \perp Y_{2t}|X_{1t}, s$

This allows us to move to the second equality above

Suppose s were observed. Then we'd have the log likelihood:

Suppose s were observed. Then we'd have the log likelihood:

$$\ell = \sum_i \sum_t \{\ell_1(Y_{1t}|X_{1t}, \alpha_1, s) + \ell_2(Y_{2t}|X_{2t}, \alpha_2, s)\}$$

Suppose s were observed. Then we'd have the log likelihood:

$$\ell = \sum_i \sum_t \{\ell_1(Y_{1t}|X_{1t}, \alpha_1, s) + \ell_2(Y_{2t}|X_{2t}, \alpha_2, s)\}$$

Two-stage estimation:

1. Estimate α_2 using only ℓ_2
2. Estimate α_1 using ℓ_1 (taking $\hat{\alpha}_2$ as given, since ℓ_1 might depend on α_2 implicitly)

Suppose s were observed. Then we'd have the log likelihood:

$$\ell = \sum_i \sum_t \{\ell_1(Y_{1t}|X_{1t}, \alpha_1, s) + \ell_2(Y_{2t}|X_{2t}, \alpha_2, s)\}$$

Two-stage estimation:

1. Estimate α_2 using only ℓ_2
2. Estimate α_1 using ℓ_1 (taking $\hat{\alpha}_2$ as given, since ℓ_1 might depend on α_2 implicitly)

Selection bias eliminated once conditioning on s and X_{1t}

But when s is unobserved, log likelihood becomes non-separable:

But when s is unobserved, log likelihood becomes non-separable:

$$\ell = \sum_i \log \left(\sum_s \pi_s \prod_t \mathcal{L}(Y_{1t}|X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s) \right)$$

But when s is unobserved, log likelihood becomes non-separable:

$$\ell = \sum_i \log \left(\sum_s \pi_s \prod_t \mathcal{L}(Y_{1t}|X_{1t}, \alpha_1, s)\mathcal{L}(Y_{2t}|X_{2t}, \alpha_2, s) \right)$$

where the π's are the population type probabilities

But when s is unobserved, log likelihood becomes non-separable:

$$\ell = \sum_i \log\left(\sum_s \pi_s \prod_t \mathcal{L}(Y_{1t}|X_{1t},\alpha_1,s)\mathcal{L}(Y_{2t}|X_{2t},\alpha_2,s)\right)$$

where the π's are the population type probabilities

Implication: Cannot estimate parts separately

Although conditional independence simplified things a bit

Resolving selection problem requires joint estimation of both components