Life Cycle Schooling and Educational Selectivity: Models and Evidence

Stephen V. Cameron and James J. Heckman
1. Models of Schooling Choices: A Brief Survey of The Literature

(1) \( P_{s-1,s}(x_s) = \Pr(D_s = 1 \mid X_s = x_s, D_{s-1} = 1) \)

(2) \( P_{s-1,s}(x_s) = \frac{\exp(x_s \beta_s)}{1 + \exp(x_s \beta_s)}. \)

\[
\text{Completed schooling} = \sum_{s=1}^{\tilde{S}} D_s
\]
\( \sum_{s=1}^{\bar{S}} D_s \mid x \) = \sum_{s=1}^{\bar{S}} s \left[ \prod_{\ell=1}^{S} P_{\ell-1,\ell}(x_\ell) \right] (1 - P_{s,s+1}(x_{s+1})) \)

2. The Effects of Heterogeneity: Are They Real - Can They Be Identified?

\( \theta = X_u \beta_u \).

**(A-1):** \( \Theta \) is independent of \( X \).

\[ \Pr(D_j = 1 \mid x, \theta) = P_{0,1}(x, \theta) \]

\[ \Pr(D_j = 1 \mid x, \theta, D_{j-1} = 1) = P_{j-1,j}(x, \theta) \]
(4) \( \Pr(S = s \mid x, \theta) = \prod_{j=1}^{S} \Pr(D_j = 1 \mid x, \theta, D_{j-1} = 1) \cdot \Pr(D_{s+1} = 0 \mid x, \theta, D_s = 1) \)

\[ \Pr(D_{S+1} = 0 \mid x, \theta, D_S = 1) = 1 \]

(5) \( \Pr(S = s \mid x) = E_{\theta}(\Pr(S = s \mid x, \theta)) \)

\[ = \int_{\theta} \Pr(S = s \mid \theta) \]
\[ x, \theta) f(\theta)d\theta \]

\[ \Pr(S \geq s \mid x) = E_\theta(\Pr(S \geq s \mid x, \theta)) \]

\[ = \int \prod_{j=1}^{S} \Pr(D_j = 1 \mid x, \theta, D_{j-1}) f(\theta)d\theta. \]
\[ \Pr(D_1 = 1 \mid x, \theta) = \frac{\exp(x\beta_1 + \theta)}{1 + \exp(x\beta_1 + \theta)}. \]

\[ \Pr(D_1 = 1 \mid x) = \frac{\exp(x\gamma_1)}{1 + \exp(x\gamma_1)}. \]

(6) \[ \text{plim}(\hat{\gamma}_1 - \beta_1) = \int_0^1 \left[ E_{X,\theta} \left( \frac{\exp(X\beta_1 + [\theta]\lambda)}{[1 + \exp(X\beta_1 + [\theta]\lambda)]^2} XX' \right) \right]^{-1} \]
\[ E_{X,\theta} \left[ \frac{\exp(X\beta_1 + [\theta]\lambda)}{[1 + \exp(X\beta_1 + [\theta]\lambda)]^2} \right] d\lambda, \]
\begin{align*}
\text{(7) } f(\theta \mid X, D_j=1, \ldots, D_1=1) = & \\
f(x) \prod_{i=1}^{j} \Pr(D_i = 1 \mid x, \theta, D_{i-1} = 1) f(\theta) d\theta \\
\frac{f(x) \prod_{i=1}^{j} \Pr(D_i = 1 \mid x, \theta, D_{i-1} = 1) f(\theta) d\theta}{f(x) \int_{\theta} \prod_{i=1}^{j} \Pr(D_i = 1 \mid x, \theta, D_{i-1} = 1) f(\theta) d\theta} \\
= & \frac{f(\theta) \prod_{i=1}^{j} \Pr(D_i = 1 \mid X, \theta, D_{i-1} = 1)}{\int_{\theta} \prod_{i=1}^{j} \Pr(D_i = 1 \mid X, \theta, D_{i-1} = 1) f(\theta) d\theta}.
\end{align*}
Observe that if as \( x \) becomes large,

\[
(D_j = 1 \mid x, \theta, D_{j-1} = 1) \text{ approaches unity},
\]

\[
\lim_{x \to \infty} f(\theta \mid x, D_j = 1, \ldots, D_1 = 1) = f(\theta).
\]

Rewrite (8) in weighted distribution form

\[
f(\theta \mid x, D_j = 1, \ldots, D_1 = 1) = f(\theta) \omega(\theta, x)
\]

\[
\omega(\theta, x) = \frac{\prod_{i=1}^{j} \Pr(D_i = 1 \mid x, \theta, D_{i-1} = 1)}{\int \prod_{i=1}^{j} \Pr(D_i = 1 \mid x, \theta, D_{i-1} = 1) f(\theta) d\theta}.
\]
3. Behavioral Models That Restrict Grade

Transition Probabilities

(A) An Ordered Discrete Choice Model

(9) $\max_s (R(s) - c(s \mid x))$.

$c(s \mid x) = c(s) \varphi(x) \varepsilon$

where

$E(\varepsilon) = 1 \text{ and } \varphi(x) \geq 0$

$R(j) - c(j \mid x) \varepsilon \geq 0$

and

9
\[ R(j) - c(j \mid x) \varepsilon \geq R(j - 1) - c(j - 1 \mid x) \varepsilon \]

and

\[ r(j) - c(j \mid x) \varepsilon \geq R(j + 1) - c(j + 1 \mid x) \varepsilon. \]

Thus a person taking \( j \) years of schooling satisfies the following inequalities:

\[
\frac{R(j) - R(j - 1)}{c(j \mid x) - c(j - 1 \mid x)} \geq \varepsilon
\]

and
\[
\frac{R(j + 1) - R(j)}{c(j + 1 \mid x) - c(j \mid x)} \leq \varepsilon
\]

and

\[
[R(j) / c(j \mid x)] \geq \varepsilon.
\]

These inequalities partition \( \varepsilon \) into intervals:

\[
(10) \quad \frac{R(j+1)-R(j)}{c(j+1 \mid x)-c(j \mid x)} < \varepsilon < \min \left[ \frac{R(j)}{c(j \mid x)}, \frac{R(j)-R(j-1)}{c(j \mid x)-c(j-1 \mid x)} \right],
\]

\[
\quad j = 0, \ldots, \bar{S}.
\]

With one additional assumption on the cost and return functions:

\[
(A-2) \quad \frac{R(j)}{R(j - 1)} < \frac{c(j \mid x)}{c(j - 1 \mid x)}
\]

which in the case of separable cost functions
\[ c(j \mid x) = c(j)\varphi(x) \] simplifies to

\[ (A-2') \quad \frac{R(j)}{R(j - 1)} < \frac{c(j)}{c(j - 1)}, \quad j = 1, \ldots, \bar{S}, \]

the region of (10) simplifies to

\[
\frac{R(j + 1) - R(j)}{[c(j + 1) - c(j)]\varphi(x)} \leq \varepsilon \leq \frac{R(j) - R(j - 1)}{[c(j) - c(j - 1)]\varphi(x)}.
\]

Letting

\[
\Pr(S = s \mid x) = \Pr \left( \frac{R(j + 1) - R(j)}{c(j + 1) - c(j)} \leq \varepsilon \leq \frac{R(j) - R(j - 1)}{c(j) - c(j - 1)} \right)
\]

and
\[ \exp(\ell(s)) = \frac{R(s + 1) - R(s)}{[c(s + 1) - c(s)]} \]

we obtain

\[ \Pr \left( \exp \frac{\ell(s)}{\varphi(x)} \leq \varepsilon \leq \exp \frac{\ell(s - 1)}{\varphi(x)} \right). \]

If, for example \( \ln \gamma \) is normal, with variance \( \sigma_{ln\varepsilon}^2 \),

\[ \Pr(S=s \mid x) = \Phi \left[ \frac{\ell(s-1) - \ln \varphi(x)}{\sigma_{ln\varepsilon}} \right] - \Phi \left[ \frac{\ell(s) - \ln \varphi(x)}{\sigma_{ln\varepsilon}} \right] \]

where \( \Phi \) is the univariate standard normal distribution. Letting

\[ \varphi(x) = \exp(-x\beta), \Pr(S = s \mid x) = \int_{\ell(s) + x\beta}^{\ell(s-1) + x\beta} dF(\ln \varepsilon). \]
Thus the probability that a person with \((s - 1)\) years of schooling completes \(s\) years of school is

\[
\Pr(D_s = 1 \mid D_{s-1} = 1, x) = \frac{\int_{\ell(s-1)+x\beta}^{\ell(s-1)+x\beta} dF(\ell n \varepsilon)}{\int_{-\infty}^{\ell(s-1)+x\beta} dF(\ell n \varepsilon)}. 
\]

(11)