Rice University
Department of Economics
Econ 515: Labor Economics, Spring 2018
Problem Set 4:

Consider the following model. Let $S_i = 0$ if an individual has a high-school degree and $S_i = 1$ if an individual has a college degree. Let $y_{i,0,t}$ and $y_{i,1,t}$, for $t = 26, 27, 28, 29$ denote the log of labor income at age t of individuals without and with a college degree, respectively. Let $X_{s,t}$ and θ_i (scalar) denote potential experience and unobserved heterogeneity, respectively. Potential experience is defined as:

$$X_{s,t} = \begin{cases} t - 18, if\ S_i = 0 \\ t - 22, if\ S_i = 1 \end{cases}$$

The unobserved components $\epsilon_{i,0,t}$ and $\epsilon_{i,1,t}$ capture, respectively, unforecastable shocks for individuals without and with a college degree. Assume that shocks are independent from $X_{s,t}$ and θ_i. Write:

$$y_{i,0,t} = \alpha_0 + \beta_0 X_{s,t} + \gamma_0 X_{s,t}^2 + \rho_0 \theta_i + \epsilon_{i,0,t}$$

$$y_{i,1,t} = \alpha_1 + \beta_1 X_{s,t} + \gamma_1 X_{s,t}^2 + 1.0\theta_i + \epsilon_{i,1,t}$$

Where $\epsilon_{i,s,t}$ is independently and normally distributed with mean zero and variance $\sigma_{s,t}^2$ for $s = 0{,}1$ and $t = 1{,}2$.

Assume that individuals maximize expected log income so an individual chooses to get a college degree if, and only if, Y_1 is greater than or equal to Y_0:

$$S = 1 \leftrightarrow E\left(\sum_{t=26}^{29}\left(y_{i,1,t} - y_{i,0,t}\right) - Z_i\delta_z - \theta_i\delta_\theta - \omega_i \middle| \mathcal{I}\right) \geq Y_0$$

Where Z_i and ω_i are, respectively, observed and unobserved non-pecuniary cost of attending college. The information set of the parents, denoted by $\mathcal{I}$, is $\mathcal{I} = \{X, Z_i, \theta_i, \omega_i\}$.

1. Estimate the model using fakedata_ps4.dta to make sure your code recovers the true parameters.
2. Estimate the model using nlsy79_homework_data.dta
3. Simulate the model and test fit.
4. Use the model to evaluate the impact of a policy that reduces tuition to zero for any student whose family income is below $15,000.
5. Estimate:
    a. Average Treatment Effect (ATE)
    b. Treatment on the Treated (TT)
    c. Local Average Treatment Effect (LATE).
6. Estimate ATE and LATE assuming that there is selection only on observables (i.e., apply matching).