Problem Set 4

Instructor : Yujung Hwang *

September 22, 2020

DUE DATE : 2020.9.30 time 11:00pm
submit your solution and code files on Blackboard page.

Consider the following income process. Log wage depends on observable characteristics Z_{it} and the unexplained component u_{it}. The unexplained income component u_{it} is composed of permanent component p_t which follows a random walk process, and a measurement error in reported log wage, m_{it}, which follows an i.i.d. normal distribution.

$$\begin{aligned} log w_{it} &= Z_{it}'\beta + u_{it} && (0.1)\\ u_{it} &= p_{it} + m_{it}, \quad m_{it} \sim N(0, \sigma_m^2) && (0.2)\\ p_{it} &= p_{it-1} + \zeta_{it}, \quad \zeta_{it} \sim N(0, \sigma_\zeta^2) && (0.3) \end{aligned}$$

Question 1. Income Process Moment Equation

Derive moment equations to estimate the following income process parameters.
(a) Assume there is no selection into work. so $P_{it} = 1$ for $\forall i, t$. Derive the moment equations to estimate the income process parameters $\sigma_\zeta^2, \sigma_m^2$.
(b) Assume there is selection into work. Set up the following auxiliary labor participation equation. The idiosyncratic preference for work η_{it} follows a joint normal distribution with permanent component shock ζ_{it}. Derive the moment equations to identify the income process parameters $\sigma_\zeta^2, \sigma_m^2$.

$$P(L_{it} = 1) = P(L_{it}^* > 0) = P(X_{it}'\gamma + \eta_{it} > 0) = P(\eta_{it} > \alpha_{it} \equiv -X_{it}'\gamma) \quad (0.4)$$

$$\begin{pmatrix} \zeta_{it} \\ \eta_{it} \end{pmatrix} \sim N(0, \begin{bmatrix} \sigma_\zeta^2 & \sigma_{\zeta\eta} \\ \sigma_{\zeta\eta} & 1 \end{bmatrix}) \quad (0.5)$$

Hint : You can use the formula for truncated bivariate standard normal distribution. That is,

$$\begin{pmatrix} X1 \\ X2 \end{pmatrix} \sim N(0, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}) \quad (0.6)$$

then

$$Var(X1|X2) = 1 - \rho^2 \quad (0.7)$$

*yujungghwang@gmail.com

$$E(X1|X2) = \rho X2 \tag{0.8}$$

$$E(X2^2|X2>c) = c\frac{\phi(c)}{(1-\Phi(c))} + 1. \tag{0.9}$$

Question 2. Estimate Income Process Using Nonlinear Least Squares

Download the dataset Pset4data.csv for this question. The dataset includes a balanced panel of 3000 individuals, including id, age, education, work status, instrumental variable for work, log wage.
Use the following participation equation.

$$P(L_{it} = 1) = P(\Pi_{it}\gamma + \eta_{it} > 0) \tag{0.10}$$

$\Pi_{it} = \{\text{Age}, \text{Age}^2, \text{i.Educ (categorical)}, \text{noveliv}\}$.
The observables for log wage includes $Z_{it} = \{\text{Age}, \text{Age}^2, \text{i.Educ (categorical)}\}$.

(a) Estimate the work participation probit regression model. Estimate the inverse mill's ratio, $\lambda(\alpha_{it})$. Report the relevant section in your code here.
(b) Write down a regression equation to estimate the unexplained income growth $\Delta u_{it} = \zeta_{it} + m_{it} - m_{it-1}$. Report your code executing the regression.
(c) Using the estimated unexplained income growth from (b), estimate the income process parameters $\sigma^2_\zeta, \sigma_{\zeta\eta}, \sigma^2_m$ using nonlinear least squares. Report the standard error by bootstrapping the whole estimation process 100 times.

[Bonus Question, +5 pt]. Compute the standard error for the estimate in (c) using the asymptotic variance formula of two-step M estimator.