London School of Economics
Department of Economics

Vassilis Hajivassiliou
Winter Term 2024

Further Topics in Econometrics (Ec485/Ec518)
Problem Set #2 – Simulation-Based Inference

1. Consider a balanced panel data set with N individuals indexed by $i = 1, \cdots, N$ each observed for $T = 5$ time periods. Consider the Multiperiod Binary Probit model defined by

$$y_{it} = \begin{cases} 1 & iff \quad x'_{it}\beta + \epsilon_{it} > 0 \\ 0 & otherwise \end{cases}$$

Suppose that the errors ϵ are fully independent from all regressor variables. They are also independent across individuals, but have the one-factor error-components structure:

$$\epsilon_{it} = \alpha_i + \nu_{it} \sim N(0, \sigma_\alpha^2 + \sigma_\nu^2)$$

where $\alpha_i \sim N(0, \sigma_\alpha^2)$ i.i.d. over i; $\nu_{it} \sim N(0, \sigma_\nu^2)$ i.i.d. over both i and ν; and α_i, ν_{it} fully mutually independent for all i and t.

Consider the 5×1 sequence of binary choices for individual i

$$y_i \equiv \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \end{pmatrix}$$

and denote the stacked vector of the sequences of the binary choices of all individuals by the $5N \times 1$ vector y. The regressors are similarly stacked into the $5N \times k$ matrix X, where k is the number of explanatory variables in the model. Our aim is to estimate the unknown parameters using the observed data set (y, X).

   (a) Can the regressor matrix X contain an intercept (vector of ones)? Why or why not?

   (b) Are the variance terms $\sigma_\alpha^2 + \sigma_\nu^2$ identified? Why or why not?

   (c) Suppose an investigator carries out Pooled Binary Probit estimation, whereby the panel data structure is completely ignored and all individual observations are treated as if they came from a single cross-section sample of dimension $5N \times 1$. Discuss the properties of this estimation approach.

(d) Describe the approach of Heckman (1981) for estimating the Random Effects Binary Probit model and explain how his approach improves on the Pooled method of (c). Discuss the order of integration necessary for implementing the Heckman estimation approach.

(e) For an individual i, characterize the vector of their conditional expectations of each binary choice given the observed data

$$E(y_i|X) \equiv \begin{pmatrix} Ey_{i1}|X \\ Ey_{i2}|X \\ Ey_{i3}|X \\ Ey_{i4}|X \\ Ey_{i5}|X \end{pmatrix}$$

Hence show how the unknown parameters of the model can be estimated through the Generalized Method of Moments approach of Avery, Hansen, and Hotz (1983). Discuss the order of integration necessary for implementing the AHH approach. Outline the properties of the AHH estimation approach.

(f) Now suppose that instead of i.i.d. over t and i, the error component ν_{it} follows an ARMA(1,1) process with autoregressive parameter γ and moving average parameter λ. How does the expression for the probability of the observed sequences of binary choices change in this case? I.e., you must characterize the probability $\Pr(y_i|X)$ and explain in which ways it differs from that probability under the one-factor error-components model of Heckman (1981). Discuss the order of integration necessary for the evaluation of $\Pr(y_i|X)$.

(g) Show how probability of the observed choices $\Pr(y_i|X)$ can be characterized through a set of linear inequality constraints on a set of correlated unobserved random variables. Use this framework to explain the Simulation-Based estimation approaches of Boersch-Supan and Hajivassiliou (1993) and Hajivassiliou and McFadden (1997). Discuss the properties of these estimation approaches.

NB: Show how your answer to this part can handle both the original error-component-only specification, as well as the error-components-plus-ARMA(1,1) error structure of part 1(g).

2. Consider the Multiperiod Autoregressive Binary Probit model defined by:

$$y_{it} = \begin{cases} 1 & iff \quad \delta y_{i,t-1} + x'_{it}\beta + \epsilon_{it} > 0 \\ 0 & otherwise \end{cases}$$

Suppose that the errors ϵ are fully independent from all regressor variables. They are also independent across individuals, but have the one-factor error-components structure:

$$\epsilon_{it} = \alpha_i + \nu_{it} \sim N(0, \sigma_\alpha^2 + \sigma_\nu^2)$$

where $\alpha_i \sim N(0, \sigma_\alpha^2)$ i.i.d. over i; $\nu_{it} \sim N(0, \sigma_\nu^2)$ i.i.d. over both i and ν; and α_i, ν_{it} fully mutually independent for all i and t.

Consider the 5×1 sequence of binary choices for individual i

$$y_i \equiv \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \\ y_{i4} \\ y_{i5} \end{pmatrix}$$

and denote the stacked vector of the sequenes of the binary choices of all individuals by the $5N \times 1$ vector y. The regressors are similarly stacked into the $5N \times k$ matrix X, where k is the number of explanatory variables in the model. Our aim is to estimate the unknown parameters using the observed data set (y, X).

(a) What additional problems are caused by the presence of the lagged dependent variable as regressor, i.e., by $\delta \neq 0$? How did Heckman (1981) propose to handle these extra problems?

(b) Explain the terms "State Dependence" and "Persistent Unobserved Heterogeneity" and discuss how these two properties can be distinguished and tested formally.

(c) Show how probability of the observed choices $\Pr(y_i|X)$ can be characterized through a set of linear inequality constraints on a set of correlated unobserved random variables. Use this framework to explain the Simulation-Based estimation approaches of Boersch-Supan and Hajivassiliou (1993) and Hajivassiliou and McFadden (1997). Discuss the properties of these estimation approaches for the more complicated error specification that, in addition to the heterogeneity term α_i, the ν_{it} error component follows an ARMA(1,1) process with autoregressive parameter γ and moving average parameter λ.

3. Consider the Simultaneous Binary Liquidity and Ordered Response Employment model of Hajivassiliou and Ioannides (2005). The model is estimated using a Panel Data set on the S_{it} and E_{it} dependent discrete variables and the matrix of explanatory factors for the two sides, X_S and X_E respectively.

   Dropping the i index for simplicity, define two latent dependent variables $y_{1t}^* \equiv S_t^*$ and $y_{2t}^* \equiv E_t^*$ that are the underpinnings of S_t and E_t according to:

   $$S_t = \begin{cases} 1 & iff \quad S_t^* > 0 \\ 0 & otherwise \end{cases} \qquad E_t = \begin{cases} -1 & iff \quad E_t^* < \theta^- \\ 0 & iff \quad \theta^- < E_t^* < \theta^+ \\ +1 & iff \quad E_t^* > \theta^+ \end{cases}$$

   Also dropping the t subscript for ease of notation, we consider the model with spillover effects on both sides:

   $$y_1^* \equiv S^* = \mathbf{1}(y_2^* < \theta^-)\delta_{01} + \mathbf{1}(y_2^* > \theta^+)\delta_{02} + x_1\beta_1 + \epsilon_1$$
   $$y_2^* \equiv E^* = \mathbf{1}(y_1^* > 0)\kappa_0 + x_2\beta_2 + \epsilon_2$$

   The contemporaneous spillover effect $\delta_0 E$ on the RHS of S^* into $\delta_{01}\mathbf{1}(\mathbf{E} = \mathbf{-1}) + \boldsymbol{\delta_{02}}\mathbf{1}(\mathbf{E} = \mathbf{1})$, i.e., into separate terms for the overemployment and the under/unemployment indicators.

   Consider the observed sequence of discrete responses S_{it} and E_{it} for individual i in period t. Stack these into their vector of choices for all periods:

   $$(S_i, E_i | X_S, X_E) \equiv \begin{pmatrix} S_{i1} \\ E_{i1} \\ S_{i2} \\ E_{i2} \\ \vdots \\ S_{i,T-1} \\ E_{i,T-1} \\ S_{iT} \\ E_{iT} \end{pmatrix}$$

   Define the joint probability of the observed discrete responses for all individuals:

   $$\Pr(S_1, E_1, S_2, E_2, \cdots, S_{N-1}, E_{N-1}, S_N, E_N | X_S, X_E)$$

   Show how this probability can be characterized through a set of linear inequality constraints on a set of correlated unobserved random variables ϵ_{1it} and ϵ_{2it} for all i and t. Use this framework to explain the Simulation-Based estimation approaches of Boersch-Supan and Hajivassiliou (1993) and Hajivassiliou and McFadden (1997). Discuss the properties of these estimation approaches for the more complicated error specifications for ϵ_{1it} and ϵ_{2it} when they include both persistent heterogeneity terms like α_i, as well as error components which follow ARMA processes.