Faculty of Economics
University of Cambridge
Lent 2020

Theory and Practice of Econometrics: II

Supervision Sheet 5

Question 1

A standard model for estimating the effects of job training or other programs on subsequent wages may be written as

$$\log(wage_{it}) = \boldsymbol{x}_{it}'\boldsymbol{\gamma} + \delta prog_{it} + \varepsilon_{it}, \tag{1}$$

where $i = 1, ..., N$ indexes individual and $t, ..., T$ indexes time periods. $wage_{it}$ denotes a measure of wages and $prog_{it}$ is a binary indicator for participation in a job training program. $\boldsymbol{x}_{it}$ is a set of observable characteristics that affect wages and may also be correlated with program participation.

The composite error term is given by $\varepsilon_{it} = \theta_t + c_i + u_{it}$. θ_t and c_i denote, respectively, unobserved time and individual effects. u_{it} represents unobserved factors that vary over i and t.

1. In evaluating the impact of the training program data is collected at two points in time and no information on $\boldsymbol{x}_{it}$ is available. At $t = 1$ $prog_{i1} = 0$ for all i. A subgroup is then chosen to participate in the program and wages are observed for the control and treatment groups in $t = 2$.

   (a) Write down an expression for $\widehat{\boldsymbol{\gamma}}$.

   (b) Using your findings from (a) or otherwise, show that for $T = 2$ the first difference and fixed effects estimators are numerically identical.

2. An analyst estimates the parameters of (1) using the fixed effects estimator. Information on $\boldsymbol{x}_{it}$ is now available.

   (a) Write down the necessary assumptions for a consistent estimator of δ.

   (b) Explain what is meant by the strict exogeneity assumption and why this might be violated in this particular case?

   (c) Should you be concerned about the stationarity properties of wages?

3. Interpret the quantity $\omega = \sigma_c^2/\sigma_c^2 + \sigma_\varepsilon^2$. For your sample you find that $\widehat{\omega} = 0.1$.

   In what sense is this statistic informative as to the choice of estimator

Question 2

A one-way error components (random effects) model has the following structure

$$y_{it} = \lambda + \boldsymbol{x}'_{it}\boldsymbol{\beta} + \alpha_i + \varepsilon_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T \tag{2}$$

Let $w_{it} = \alpha_i + \varepsilon_{it}$ represent the sum of the individual specific random component and an idiosyncratic error varying over both dimensions. The distributional assumptions, in matrix form, are given by, $\boldsymbol{\alpha} \sim (\mathbf{0}, \sigma^2_{\boldsymbol{\alpha}} \boldsymbol{I}_N)$, $\boldsymbol{\varepsilon} \sim (\mathbf{0}, \sigma^2_{\varepsilon} \boldsymbol{I}_{NT})$, and $\boldsymbol{\alpha}$, $\boldsymbol{\varepsilon}$ independent.

1. Write down the covariance matrix for $\boldsymbol{w}_i = \alpha_i \boldsymbol{i}_T + \boldsymbol{\varepsilon}_i$, where $\boldsymbol{w}_i$ and $\boldsymbol{\varepsilon}_i$ are $T \times 1$ vectors, and $\boldsymbol{i}_T$ is a $T \times 1$ unit vector.
2. In what sense is the Random Effects (RE) estimator a Generalised Least Squares estimator?
3. Discuss how both the OLS and FE estimators may be considered as (different) limiting cases of the RE estimator.

Question 3

Consider the following linear regression for panel data including a fixed effect, η_i, with N individuals and T time periods:

$$y_{it} = \alpha y_{it-1} + \beta_0 x_{it} + \beta_1 x_{it-1} + \eta_i + v_{it} \quad (i = 1, \ldots, N; t = 2, \ldots, T)$$

Discuss the identification and estimation of the parameters of a model of this type when T is small and N is large, given

$$E(v_{it} | x_{i1}, ..., x_{iT}, \eta_i) = 0$$

You are also told that the errors are serially correlated.

Question 4

Let $\boldsymbol{z}_1$ be a vector of variables, let z_2 be a continuous variable and let d_1 be a dummy variable.

a. In the model

$$\Pr(y = 1|\boldsymbol{z}_1, z_2) = \boldsymbol{\Phi}(\boldsymbol{z}_1\boldsymbol{\delta}_1 + \gamma_1 z_2 + \gamma_2 z_2^2),$$

find the partial effect of z_2 on the response probability. How would you estimate this partial effect?

b. In the model

$$\Pr(y = 1|\boldsymbol{z}_1, z_2) = \boldsymbol{\Phi}(\boldsymbol{z}_1\boldsymbol{\delta}_1 + \gamma_1 z_2 + \gamma_2 d_1 + \gamma_3 z_2 d_1),$$

find the partial effect of z_2. How would you measure the effect of d_1 on the response probability? How would you estimate these effects?

M. Weeks

Lent 2020