CERGE-EI
Summer 2013/2014
Instructors: Nikolas Mittag, Dragana Stanišić
TA: Jelena Plazonja, Gega Todua
Date: 20/05/2014

ECONOMETRICS II

Exercise session #4
(Panel Data)

Problem 1. Consider a simple model to estimate the effect of tablet ownership on college grade point average for graduating seniors at a large public university:

$$GPA = \beta_0 + \beta_1 tablet + u$$

where *tablet* is a binary variable indicating tablet ownership.

(i) Why might tablet ownership be correlated with u?

(ii) Explain why *tablet* is likely to be related to parents' annual income. Does this mean parental income is a good IV for *tablet*? Why or why not?

(iii) Suppose that, four years ago, the university gave grants to buy tablets to roughly one half of the incoming students, and the students who received grants were randomly chosen. Carefully explain how you would use this information to construct an instrumental variable for *tablet*.

Problem 2. In order to determine the effects of collegiate athletic performance on applicants, you collect data on applications for a sample of Division I colleges for 2000, 2005 and 2010.

(i) What measure of athletic success would you include in an equation?

(ii) What other factors might you control for in the equation?

(iii) Write an equation that allows you to estimate the effects of athletic success on the percentage change in applications. How would you estimate this equation? Why would you choose this method?

Problem 3. In a random effects model, define the composite error $v_{it} = a_i + u_{it}$, where a_i is uncorrelated with u_{it}, and the u_{it} errors have variance σ^2 and are serially uncorrelated. Define $e_{it} = v_{it} - \lambda \bar{v}_i$ where $\lambda = 1 - \sqrt{\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_a^2}}$, $\bar{v}_i = \frac{1}{T}\sum_{i=1}^{T} v_{it}$. Show that the error e_{it} have mean zero, constant variance, and are serially uncorrelated.

Problem 4. Suppose we want to estimate the effect of several variables on annual saving and that we have a panel data set on individuals collected on January 31^{st} 1990 and January 31^{st} 1992. If we include a year dummy for 1992 and use first differencing, can we also include age in the original model? Explain.

Problem 5. Consider the following two period fixed effect model

$$y_{it} = \beta_0 + \beta_1 Dummy2_t + \beta_2 x_{it} + a_i + \epsilon_{it}$$

where i denotes cross section units, $Dummy2$ indicates second period.

(i) Does the sign of the β_2 in this equation affect the direction of bias in the OLS estimation in the FD equation if $CORR(\Delta x_i, \Delta\epsilon_i) < 0$?

(ii) List the key assumptions for validity of OLS estimation in FD.

Problem 6. Suppose that the idiosyncratic error u_{it} in

$$y_{it} = \beta_1 x_{it}^1 + \beta_2 x_{it}^2 + \ldots + \beta_k x_{it}^k + a_i + u_{it},\ t = 1, 2, \ldots, T$$

are serially uncorrelated with constant variance, σ_u^2. Show that the correlation between adjacent differences, Δu_{it} and $\Delta u_{i,t+1}$ is -0.5. Therefore, under the ideal FE assumptions, first differencing induces negative serial correlation of a known value.