Temple University
Economics Department

Econometrics

The data for this problem set are in REMDATA.TXT on this web site. The first column of data is the observations on x_{it}. There are ten observations, t=1,2,…,10, for each of five persons,i=1,2,…5. The column of data immediately following x_{it} is the data for the dependent variable. It s organized the same way: t=1,2,…,10 periods for each of i=1,2,…,5 persons.

1. The data are believed to have been generated by the model shown below.

$$y_{it} = \beta_1 + \beta_2 x_{it} + \varepsilon_{it}$$

$$i = 1,2,3,4,5 \quad t = 1,2,3,\ldots,10$$

The data are a cross section of time series. There are five persons and ten time periods. The independent variable is an index for years of experience. The dependent variable is an index for hourly earnings.

a. Estimate the parameters of the model by ordinary least squares.

b. Estimate the parameters of the model with a set of dummy variables to shift the intercept for each individual.

c. Estimate the parameters with a set of dummy variables to shift the intercept for each time period.

d. Estimate the model with both the individual and the time dummies.

e. Which specification among the OLS and LSDV models do you prefer and why?

f. Estimate the model parameters allowing for random effects across individuals only.

g. Estimate the model parameters allowing for random effects across time periods only.

h. Estimate the model parameters allowing for both time and individual random effects.

i. Among the random effects models, which specification do you prefer and why?

j. Now consider the OLS model, your preferred LSDV model, and your preferred REM model. Which of these three do you prefer and why?