Panel Data Analysis

Homework 3
Prf. José Fajardo

Deadline: February 12^{th}, 2019 at 14:00am

1. Using two datasets time_var.dta and time_invar.dta, consider the following wage equation:

   $$lwage_{it} = \alpha_0 + \alpha_1 ability_i + \alpha_2 medu_i + \alpha_3 fedu_i + \alpha_4 d_i + \alpha_5 siblings_i + \beta_1 ed_{it} + \beta_2 pexp_{it} + \varepsilon_{it}$$

   Notice that all the α coefficients are associated with time-invariant cross section data, while β are with time-variant panel data series.

   a) Formulate, estimate, and compare the pooled or population-averaged based on OLS and OLS with panel-robust standard errors, respectively. In addition to pooled model, three different variable transformations should be considered and compared: (1) first-difference, (2) between (or group means), and (3) within (or deviations from group means). Note: not all coefficients can be estimated for all models. Why?

   b) Formulate, estimate, and compare the fixed-effects and random-effects panel data models based on OLS and OLS with panel-robust standard errors, respectively. Setup and perform hypothesis testings to choose a proper panel data model: (1) pool or not to pool? (2) fixed-effects or random-effects?

      Read: Koop, G.M. and Tobias, J. (2004) Learning about heterogeneity in returns to schooling. Journal of Applied Econometrics, 19 (7). pp. 827-849. and
      Joshua C. C. Chan, "Replication of the Results in 'Learning about Heterogeneity in Returns to Schooling'", Journal of Applied Econometrics, Vol. 20. No. 3, 2005, pp. 439-443.

2. The data in the file produtivity.txt is a panel on the following variables for the 48 states, 17 years,

   The basic model of interest is

   $$Y_{it} = \beta_1 X1_{it} + \beta_2 X2_{it} + \beta_3 X3_{it} + \beta_4 X4_{it} + \beta_5 X5_{it} + c_i + \varepsilon_{it}$$

   Where Y is $logGSP$, $X1$ is $logPC$, $X2$ is $logHWY$, $X3$ is $logWATER$, $X4$ is $logUTIL$ and $X5$ is $logEMP$.

   This is a Cobb-Douglas production function.

STATE	= state name
YR	= 1970,...,1986
P_{CAP}	= public capital
HWY	= highway capital
WATER	= water utility capital
UTIL	= utility capital
PC	= private capital
GSP	= gross state product
EMP	= employment
UNEMP	= unemployment rate

a) Fit the pooled model and report your results

b) Fit a random effects model and a fixed effects model. Use your model results to decide which one is the preferred model. If you find that neither panel data model is preferred to the pooled model, show how you reached that conclusion

Read: Munnell, A. "Why Has Productivity Declined? Productivity and Public Investment." New England Economic Review, 1990, 3-22.