Econometrics-2022-2023.
Home assignment 10. Panel Data Models
To be submitted by March 26, 23:55

1. [30 marks] The following estimates were made as part of a study of whether countries with low income per head had higher growth rates than countries with higher income per head (convergence). Data was obtained over a period of four years on 121 countries.
The first estimates are pooled ordinary least squares estimates:

$$gr_{it} = \underset{(0.358)}{-0.129} - \underset{(0.004)}{0.023}\, pgdp_{it} + \underset{(0.018)}{0.139} inv_{it} + e_{1it} \quad R^2 = 0.133 \quad ; \quad . i = 1,..., N,\ t = 1,.., T \quad (1)$$

The second estimates are fixed effects (least squares dummy variable variant):

$$gr_{it} = \underset{(0.358)}{3.524} - \underset{(0.012)}{0.046}\, pgdp_{it} - \underset{(0.032)}{0.020} inv_{it} + e_{2it} \quad R^2 = 0.541 \quad ; \ i = 1, 2, ..., N \ \ t = 1, 2, ..., T \qquad (2)$$

$$N = 121,\ T = 4, \quad .$$

Standard errors are in brackets. *gr* is the average percentage annual rate of growth of *gdp* per head, *pgdp* is the *gdp* per head in constant US dollar, *inv* is the average percentage ratio of investment to *gdp*. e_1 and e_2 are the ordinary least squares residuals. R^2 is the conventionally calculated coefficient of determination.

1.1. [10 marks] Explain theoretically what are the difference between two models under consideration: pooled regression and least squares dummy variable fixed effect method. What are relative advantages and disadvantages of each type of model?

1.2. [10 marks] Comment on the differences in R^2 for each equation. Equation (1) is restricted version of equation (2). Test the validity of the restrictions. From these results, what would you conclude about the convergence of *gdp* per head? Give details.

1.3. [10 marks] At the seminar where this research was discussed one of the participants proposed to use some other fixed effect methods like the within-groups and first differences arguing that they have many advantages over LSDV. Explain the differences between these methods and comment on the proposal of the participant.

2. [20 marks] A student was told to take the data for 26 OECD countries for three years (2011-2013), on w_{it} (average annual percentage rate of growth of wages for country *i* during time period *t* ($t = 1$ for 2011, $t = 2$ for 2012, $t = 3$ for 2013) and p_{it} (average annual percentage rate of growth of productivity for country *i* during time period *t*). He found all data for 2011 and 2013 but failed to find any data for 2012. Being limited in time he decided to use the mean values of 2011 and 2013 as the interpolations for 2012, namely $w_{i2} = \frac{w_{i1} + w_{i3}}{2}$, $p_{i2} = \frac{p_{i1} + p_{i3}}{2}$, $i = 1, ...,26$. Then he estimates Fixed Effect regression $w_{it} = \beta_1 + \beta_2 p_{it} + \alpha_i + u_{it}$ $i = 1, ...,26;$ $t = 1, 2, 3.$ Here α_i is unobserved heterogeneity term that in Fixed Effect regression supposed to be a combined result of influence of some non-random factors.

2.1. [10 marks] What are advantages in analysis of panel data comparing to cross-section regression and time series analysis? Explain briefly what is the Fixed Effect approach to the estimation of the Panel Model. Why here it is the only possible approach?

2.2. [10 marks] What are comparative advantages and disadvantages of Within group and First difference methods in estimating panel regressions? Why the application of any of these methods leads to autocorrelation problem? How serious is this problem in each of these methods? How it could be mitigated?

In answering the practical questions, no general theoretical/mathematical explanations are acceptable, nor any information on the topic that is not directly related to the question.

3. [50 marks] Using the file ***nondurables.wf1*** investigate the factors influencing everyday expenditures of Americans applying panel data models. File contains data on expenditures on such nondurables as CLOT, FLOW, FOOD, GASO, MAGS, TOB, TOYS (see description of variables in file ***HA10 data description.pdf***).

3.1. [10 marks] Estimate pooled regression to evaluate income and relative prices elasticities of expenditure on nondurables. Briefly describe your result. Compare pooled regression with some separate regressions results: what are their relative advantages and disadvantages? What test is appropriate to support your conclusion?

3.2. [10 marks] Use LSDV method to estimate fixed effect panel data model. Interpret your results both economically and graphically.

3.3. [10 marks] Test whether LSDV model is significantly better as pooled regression. Do both manual and automatic tests and compare results.

3.4. [10 marks] Use first difference fixed effect method to estimate your panel data model. What are its advantages and disadvantages. Explain what are relative advantages and disadvantages of Within Group method (no need to do it for your data as it can be shown two methods LSDV and Within Group are equivalent in estimation of the model). What information can be obtained only from LSDV method?

3.5. [10 marks] Use random effect method to estimate the model. Compare results of estimation of the model using fixed effect and random effect methods.

The coefficients obtained by two different methods are very close to each other. So if you attempt to do Durbin-Wu-Hausman test the results will be disappointing: the message appears

* Cross-section test variance is invalid. Hausman statistic set to zero.

The reason is that when there are small differences in coefficients the covariance matrix may not be positive definite. In fact when the results of the two methods are so close to each other there is no need for special test to come to the conclusion. What is your conclusion? What is economic meaning of your conclusion?

Try to change specification and remove logarithm of income per capita from the model. Compare again fixed and random effect methods and do Durbin-Wu-Hausman test. What are the assumptions, decision rule and conclusion from the test? What are advantages of choosing proper model and potential risks of the improper choice?