Advanced Applied Econometrics
OECD, December 2010
Thijs van Rens

Workshop #3

Exercise 1. Dynamic Panel Data models For this exercise, we will use cross-country panel data on educational attainment and GDP, available from my website at www.crei.cat/~vanrens/educ (the same data we use for the workshop on DD). The aim is to estimate the social return to education, allowing for dynamic effects. The estimation equation is given by,

$$\log Y_{it} = \alpha_i + \beta_0 S_{it} + \gamma \log Y_{it-1} + \beta_1 S_{it-1} + \varepsilon_{it}$$

where Y_{it} is GDP per worker in country i and year t and S_{it} is educational attainment (in years). Unless stated otherwise, we treat education as strictly exogenous.

1. Estimate the model using pooled OLS (ignoring the fixed effects). The estimates of this regression are not consistent. In which direction do we expect the bias to go for γ? And for β_0 and β_1?
2. Estimate the model using the within-group (FE) estimator. The estimates from this regression are not consistent either. In which direction do we expect the bias to go for γ? And for β_0 and β_1?
3. Use the estimates in parts 1 and 2 to find an upper and a lower bound for the short run return to education.
4. Estimate the model using the Anderson-Hsiao instrumental variables estimator. Under some conditions, this estimator is consistent. Are the estimates consistent with the bounds from part 3?
5. The Anderson-Hsiao estimator may be biased because of weak instruments. Check whether this is an issue in this case.
6. Estimate the model using the Arellano-Bond GMM estimator. Under some conditions, this estimator is consistent. Are the estimates consistent with the bounds from part 3?
7. The Arellano-Bond estimator may be biased because of weak instruments. Is this more or less likely to be an issue than for the Anderson-Hsiao estimator? Check whether this is an issue in this case.
8. Test for autocorrelation in the error term. What is the purpose of this test? What do you conclude?
9. Estimate the model using the Blundell-Bond system GMM estimator. Under some conditions, this estimator is consistent. Are the estimates consistent with the bounds from part 3?
10. The system GMM estimator may be biased because of weak instruments. Is this more or less likely to be an issue than for the Arellano-Bond estimator? Check whether this is an issue in this case. What other (dis)advantages does the system GMM estimator have over the Arellano-Bond estimator.
11. Test for autocorrelation in the error term. What is the purpose of this test? What do you conclude?
12. So far, we have assumed education is exogenous. Why is this assumption likely to be violated? Now, assume education is not exogenous, but predetermined. Answer again questions 4-11.

13. Using all estimates above and your best judgment, what is the short run social return to education?

14. Based on your favorite specification, what is the long run social return to education? Using the delta-method, calculate the standard errors on the long run return. Is the long run return more or less robust across specifications than the short run return? Why?

Exercise 2. Standard errors In each of the following regressions, what is your main concern with the usual OLS standard errors and how would you solve it? How would you implement your solution in Stata?

1. To assess the effect of human capital on growth, you estimate the following regression using cross-country panel data,

$$y_{it} = \beta_0 + \beta_1 S_{it} + \varepsilon_{it}$$

where y_{it} is GDP of country i in year t and S_{it} is the average level of schooling in that country. You may assume that $E[\varepsilon_{it} S_{it}] = 0$.

2. You are interested in estimating the effect of financial development on growth. Using cross-country panel data, you run a regression of the following form,

$$y_{jt} = \beta_0 + \beta_1 FD_{jt} + x'_{jt}\beta_2 + \varepsilon_{jt}$$

where y_{jt} is GDP of country j in year t and FD_{jt} is an index of financial development. We assume $E[\varepsilon_{jt} FD_{jt}] = E[\varepsilon_{jt} x_{jt}] = 0$.

3. In the model from the previous part, you want to control for time-invariant unobservable differences across countries, and therefore include country dummies in the regression.

4. Arellano, Bai and Zhang (2008) argue that small firms grow faster because they face more severe borrowing constraints. Thus, we would expect the growth differential between small and large firms to be smaller in countries with more developed financial markets. They run the following OLS regression,

$$g_{jc} = \beta_0 + \beta_2 s_{jc} + \beta_2 s_{jc} * FD_c + \varepsilon_{jc}$$

where g_{jc} is the growth rate of firm j in country c, s_{jc} is the size of that firm and FD_c is an index of financial development of the country in which that firm is located, and test whether $\beta_2 \neq 0$.

5. You are interested in the effect of exporting on the productivity of firms in Argentina. You run the following regression,

$$p_{ijt} = \beta_0 + \beta_1 E_{ijt} + \beta_2 C_{jt} + \gamma_i + \delta_t + \varepsilon_{ijt}$$

where p_{ijt} is productivity of firm i in sector j in year t, E_{ijt} is a dummy that takes value 1 if the firm exports, C_{jt} is a control variable, which measures the competitiveness of an industry, and γ_i are the coefficients on firm dummies (fixed effects). The problem is, we do not observe C_{jt} and no Argentinian data are available to estimate it. However, we can use US data to estimate C_{jt} and use those estimates in the regression for Argentinian firms, assuming that industry competitiveness is comparable across countries.