

## *Exercise Session 2*

### **Problem 1**

There is a controversial debate whether and to what extent the right to carry a gun influences crime. Proponents of so-called “Carrying a Concealed Weapon” (CCW) laws argue that the deterrent effect of guns prevents crime, whereas opponents argue that the public availability of guns increases their usage and thus makes it easier to commit crimes. In the following exercise we will empirically investigate this topic. We will work with the dataset “Guns”, a balanced panel containing observations on criminal and demographic variables for all US states and the years 1977-1999. The dataset comes with the package AER.

Consider the model below:

$$\begin{aligned} \log(violent_{it}) = & \beta_0 + \beta_1 law_{it} + \beta_2 density_{it} + \beta_3 income_{it} + \beta_4 population_{it} + \\ & + \beta_5 afam_{it} + \beta_6 cauc_{it} + \beta_7 male_{it} + u_{it} \quad (1) \end{aligned}$$

1. Verify that Guns is a balanced panel.
2. Estimate equation (1) by pooled OLS. Discuss the issues your estimates might have. What assumption on the error term should we impose to obtain consistent estimates?
3. Re-estimate the model with standard errors clustered at a state level. What they assume and what they allow for? Compare the clustered standard errors and the ones robust only to heteroskedasticity.
4. Discuss variables that vary across states but plausibly vary little, or not at all, over time and that could cause omitted variable bias in regression of part 2.
5. Suppose we run one-way FE regression now. What is the underlying assumption on the error term in this case? Explain how it overcomes the problems in part 2 and 4. Verify that LSDV and FE methods produce the same coefficient estimates. Test whether the fixed effects are jointly significant.

6. If we were to add some State-level geographic characteristics (on top of the FEs) in our model as explanatory variables, would we be able to identify the coefficients on them? Explain your answer.
7. Explain why you would need to add time fixed effects. Estimate the equation by two-way fixed effects and test whether both state and time fixed effects are jointly significant.
8. Now assume the model is RE. List the assumptions of RE model, explain what is the difference with Pooled OLS in terms of coefficient estimates. Perform the estimation manually, and then use a direct command.
9. If you were to choose between random effects and OLS, explain which formal test you would perform and perform it.
10. Compare models from part 2, 5, 7 and 8. Comment.

## Problem 2

Consider the following panel data model:

$$Y_{it} = X'_{it}\beta + u_i + \epsilon_{it}, \quad E[X_{is}\epsilon_{it}] = 0 \quad \forall s = 1, \dots, T$$

1. Discuss finite sample properties of the within estimator
2. Derive the asymptotic variance of the within estimator and discuss a suitable estimator for it.