School of Public Health PHPM110062

Mengcen Qian Spring 2017

**Problem Set 6**

**Due on 05/19**

The following problem set question comes from your slides, “Panel data models: methods”.

Suppose that you are interested in the relationship between the death penalty and the murder rate in the US, and you collect data on the murder rate in the 50 states for 1990 and 1993. Results are summarized in the following table.

|  |  |  |
| --- | --- | --- |
| y = murders per 100,000 people | (1) | (2) |
|  |  |  |
| Executions | 0.90\*\*\* | -0.10\*\* |
|  | (0.220) | (0.040) |
|  |  |  |
| Observations | 100 | 100 |
| State FE | No | Yes |

1. What type of dataset are we using in the above analysis? (*Hint: time series, panel data, or cross sectional data*)
2. What might be the explanation for the unexpected positive impact of executions in col.(1)? (*Hint: We’ve covered several model issues, heteroskedasticity, endogeneity (omitted variable bias), serial correlation, collinearity, and irrelevant variable. Which one do you think could be the best explanation?*)
3. Assume that we are not able to fix Model (1) by adding variables (observables). Instead, we include state fixed effects in the regression. That is, we fit a fixed-effect model to the data. Use your own words to clarify how and why state FE would help us? (*Hint: what does state FE capture? With state FE, what kind of variations in the variable ‘Executions’ contribute to the identification of its impact on murder rate?*)