Chapter 13

The student interpreted these results as suggesting that, all other variables constant, every 1 million additional citizens is associated with a 0.039 increase in the total number of medals won, while being a host country is associated with an additional 14.79 medals won in the country that hosted the games. More importantly, the student focused on the fact that GDP per capita became statistically insignificant and the estimated marginal effect of being a host country shrank by roughly two-thirds when estimating the fixed-effects model rather than the pooled cross-section model. Considering these changes in more detail, the student decided that they were not that surprising given that the countries winning the most medals in the four Olympic Games included in the sample were the United States, Russia, and China and that of those countries, the United States (1996) and China (2008) both hosted the Olympics during the time periods covered by her panel.

Based on this work, from the initial idea through data collection, empirical analysis, interpretation of results, and effective write-up, our former student was judged to have conducted a successful empirical research project.

LOOKING AHEAD TO CHAPTER 14

We occasionally encounter situations in which, due to certain aspects of the data we collect, we are unable to uniquely identify the marginal effect that we wish to estimate. For instance, we have difficulty estimating the marginal effect of obesity on health because obese people are not only obese but they are also more likely to have high blood pressure, less likely to exercise regularly, etc. Accordingly, if we simply regress overall health on an indicator of obesity, we confound the effects that obesity, high blood pressure, lack of exercise, and so on, have on health, thereby rendering our estimated marginal effect of obesity on health biased. Fortunately, there are empirical tools that we can use to help us better identify the estimated marginal effects in such cases. We discuss these tools in Chapter 14.

Problems

13.1 You have a data set that has two years of individual data on wages, experience, education level, and score on a job proficiency test. Call the two time periods time period 1 and time period 2.

a. Write out the model assuming you organize the data set as a pooled cross-section.

b. Write out a new model that allows the intercept to change between the two years.

c. Write out another model that allows both the slope and the intercept to change between the two years.

d. Describe in detail how you would test if the slopes differ between the two time periods.

13.2 As in problem 13.1, you have a data set that has two years of individual data on wages, experience, education level, and score on a job proficiency test. Now you are going to exploit the time nature of these data.

a. Describe how you would estimate this model through first-differencing.

b. Describe how you would estimate this model through fixed effects by demeaning the observations.

c. Describe how you would estimate this model through fixed effects using dummy variables.

d. Will you get different results if you use first-differencing or fixed effects?

e. Why would you use panel data techniques over pooled cross-section?

f. What will happen to the education variable when you use either first-differencing or fixed effects?

13.3 Explain why you can't use an independent cross-section to use panel data techniques such as first-differencing and fixed effects.

Exercises

E13.1 Use the file Life Expectancy.xls.

a. Regress Life Expectancy on Population, GDP, and Health Expenditure as a pooled cross-section.

b. Regress Life Expectancy on Population, GDP, and Health Expenditure as a pooled cross-section controlling for the number of years.

c. Now regress Life Expectancy on GDP per Capita and Health Expenditure as a pooled cross-section, controlling for the number of years, and compare your results to what you found in parts (a) and (b). Which specification do you feel is a better fit?

E13.2 Use the file Life Expectancy.xls.

a. Use first differencing to estimate the regression of Life Expectancy on Population, GDP, and Health Expenditure.

b. Now use first-differencing to estimate the regression of Life Expectancy on GDP per Capita and Health Expenditure.

c. Compare your results to what you found in problem C13.1.

E13.3 Estimate a fixed-effects model using the NFL data with only using years 2007 and 2009. Comment on the results, and compare your estimates to the first-differenced estimates in Excel Regression Output 13.3.

E13.4 Estimate a random-effects model using the Olympics data. Comment on the results. Do you think that this is the best method to estimate this model? Explain.

E13.5 Use the file Life Expectancy.xls.

a. Estimate the model Life Expectancy on Population, GDP, and Health Expenditure using the dummy variable method of fixed effects.

b. Estimate the model Life Expectancy on Population, GDP, and Health Expenditure using the differencing method of fixed effects. Make sure your results match in parts (a) and (b).

c. Estimate the model Life Expectancy on Population, GDP, and Health Expenditure using random effects.

d. Comment on the differences.