

### Practice Set 4: Panel Data Methods

1. Kiel and McClain (1995) studied the effect that a new garbage incinerator had on housing values in North Andover, Massachusetts. The rumors that a new incinerator would be built in North Andover began after 1978, and construction began in 1981. We use data on prices of houses that sold in 1978 and another sample on those that sold in 1981. Use the data in KIELMC.dta for this exercise. Consider the model

$$\log(\text{price}) = \beta_0 + \delta_0 y81 + \beta_1 \log(\text{dist}) + \delta_1 y81 \log(\text{dist}) + u \quad (1)$$

where

$\text{price}$  = house price  
 $y81$  = year dummy for 1981  
 $\text{dist}$  = distance from each home to the incinerator site, in feet

- (a) Estimate the model using OLS. Interpret all slopes in the regression.
  - (b) Add  $\text{age}$ ,  $\text{age}^2$ ,  $\text{rooms}$ ,  $\text{baths}$ ,  $\log(\text{intst})$ ,  $\log(\text{land})$ , and  $\log(\text{area})$ . What do you conclude about the effect of the distance to the incinerator site in 1978? And about the effect of the incinerator on housing values?
  - (c) Compute the p-value of the null hypothesis that the incinerator does not affect house prices against the one sided alternative.
2. Use the state-level data on murder rates and executions in MURDER.dta for the following exercise. Consider the unobserved effects model

$$\text{mrd rte}_{it} = \theta_t + \beta_1 \text{exec}_{it} + \beta_2 \text{unem}_{it} + a_i + u_{it} \quad (2)$$

where  $\theta_t$  simply denotes different year intercepts,  $a_i$  is the unobserved state effect.

$\text{mrd rte}$  = murders per 100,000 people  
 $\text{exec}$  = total executions in the past three years  
 $\text{unem}$  = annual unemployment rate

- (a) If past executions of convicted murderers have a deterrent effect, what should be the sign of  $\beta_1$ ? What sign do you think  $\beta_2$  should have?
  - (b) Using just the years 1990 and 1993, estimate the equation by pooled OLS. Do you find any evidence for a deterrent effect?
  - (c) Now, using 1990 and 1993, estimate the equation in first differences. Is there evidence of a deterrent effect? Answer using both the usual standard errors and the heteroskedastic-robust standard errors.
  - (d) Check that STATA fixed effects for these two years gives the same result as in (c).
  - (e) Drop Texas and repeat (c). What is going on?
  - (f) Use all three years and estimate the model by Pool OLS and fixed effects. Discuss the size and statistical significance of the deterrent effect.
3. Use the data file ecpfmas.dta for this exercise.

Consider estimation of the following model

$$\ln q_{it} = \beta_1 + \beta_2 \ln x_{it} + z'_{it} \delta + \eta_i + \varepsilon_{it},$$

where  $q_{it}$  is food expenditure,  $x_{it}$  is total expenditure, and  $z_{it}$  is a vector of household characteristics.

- a. Estimate the model by OLS, including among the explanatory variables a dummy indicating if the wife works, number of children, number of adults, age of the husband, and dummies for the education of the husband and for the interview week.
- b. Estimate the model by OLS in first differences and by using the Within Groups estimator.

ecpfmas.dta

Data for a subsample of the Encuesta Continua de Presupuestos Familiares 1985-1995. The variables included in the file are the following:

1. maxcc - Number of quarters that the household is included in the sample.
2. ind - Household identifier
3. cc - Number of the interview
4. year - Year
5. week - Week
6. hedad - Age of the husband
7. medad - Age of the wife
8. mmt - Female labor force participation (1: employed, 2: unemployed; 3: out of the labor force)
9. heduc - Education of the husband (1: without studies; 2: Primary education; 3: secondary education; 4: University education; 5: more than University)
10. hijos - Number of children younger than 18
11. adultos - Number of adults in the household
12. galimr - Food expenditure
13. gtotalr - Total expenditure
14. rentar - Household income