Problem set 4

Due on Thursday November 3, 2022 (by 11:59 PM EST)

Note: No credit will be given if you report only the final answers without showing formulas and calculations when appropriate. This applies to both theoretical and empirical questions. For the empirical questions, make sure to upload the R scripts and output on Latte. No credit will be given if the R output is missing.

Problem 1

Two authors published a study in 1992 of the effect of minimum wages on teenage employment using a U.S. state panel. The paper used annual observations for the years 1977-1989 and included all 50 states plus the District of Columbia. The estimated equation is of the following type

$$E_{it} = \beta_0 + \beta_1 (M_{it}/W_{it}) + \gamma_2 D2_i + ... + \gamma_n D51_i + \delta_2 B2_t + ... + \delta_T B13_t + u_{it},$$

where E is the employment to population ratio of teenagers, M is the nominal minimum wage, and W is average wage in the state. In addition, other explanatory variables, such as the prime-age male unemployment rate, and the teenage population share were included.

(a) Estimating the model by OLS but **including only time fixed effects** results in the following output

$$\underset{}{\hat{E}_{it}} = \underset{}{\hat{\beta}_0} \underset{(0.08)}{- 0.33 \times (M_{it}/W_{it})} + \underset{(0.28)}{0.35(SHY_{it})} \underset{(0.13)}{- 1.53 \times uram_{it}}; \qquad \overline{R}^2 = 0.20$$

where SHY is the proportion of teenagers in the population, and $uram$ is the prime-age male unemployment rate. Coefficients for the time fixed effects are not reported. Numbers in parenthesis are clustered standard errors by state. Are the coefficients statistically significant?

(b) **Adding state fixed effects** changed the above equation as follows:

$$\underset{}{\hat{E}_{it}} = \underset{}{\hat{\beta}_0} \underset{(0.10)}{+ 0.07 \times (M_{it}/W_{it})} \underset{(0.22)}{- 0.19 \times (SHY_{it})} \underset{(0.11)}{- 0.54 \times uram_{it}}; \qquad \overline{R}^2 = 0.69$$

Compare the two results. Why would the inclusion of state fixed effects change the coefficients in this way?

(c) Interestingly, the significance of each coefficient decreased, yet $\overline{R}^2$ increased. What does this result tell you about testing the hypothesis that all of the state fixed effects can be removed from the model? How would you test for such a hypothesis?

Problem 2

In 1979 the town of North Andover, Mass. announced it would build a waste treatment plant. Here we investigate the effect of the waste treatment plant on housing values in North Andover, using data on prices of houses sold in 1978 and another sample of houses sold in 1981, after the plant was completed. Let rprice denote the house price in real terms, i.e. adjusted for inflation.

We begin by estimating a simple model using only data for 1981 (n = 142):

$\text{rprice}_i = 120{,}700 - 30{,}100\ \text{nearplant}_i$

(3,090) (5,830)

where nearplant_i is a binary variable equal to one if the house is near the waste treatment plant, and zero otherwise.

a. What is the price of homes not near the treatment plant? What is the price of homes near the treatment plant?

b. To calculate the effect of a waste treatment plant on housing prices we estimate the following model using data from the two period, where the dependent variable is the logarithm of real house price and the sample size is 321

$\log(\text{rprice}_{it}) = 11.3 + .457\ \text{year81}_{it} - .340\ \text{nearplant}_{it} - 0.063\ \text{year81}_{it} \times \text{nearplant}_{it}$

(.31) (.045) (.055) (.023)

where year81 = 1 if the year is 1981, = 0 if the year is 1978.

What is the interpretation of the coefficient on year81 (be specific about what this variable controls for)? What is the interpretation of the coefficient on nearplant (be specific about what this variable controls for)?

c. According to the regression in part (b), what is the effect of the waste treatment plant on house prices? Can this be interpreted as a causal effect? Explain.

Problem 3

Arora and Vamvakidis (2005) examine the extent to which South African economic growth is an engine of growth in sub-Saharan Africa using the following regression model which includes two interaction terms, $(SAF)_t(TRADE)_{it}$ and $(SAF)_t(DIST)_{it}$.

$$\%\Delta GDP_{it} = \beta_0 + \beta_1 (SAF)_t + \beta_2 (SAF)_t (TRADE)_{it} + \beta_3 (SAF)_t (DIST)_{it} + \beta_4 X_{4,it} + \cdots + \beta_k X_{k,it} + (country\ \ dummies) + u_{it}$$

where $\%\Delta GDP_{it}$ is the per capita real GDP growth rate in country $i\,(=1,\cdots,47)$ in year $t\,(=1960\sim1999)$, SAF_t the per capita real GDP growth rate in South Africa in year *t*, $TRADE_{it}$ the share of exports to South Africa in country *i*'s total exports in year *t*, and $DIST_{it}$ the distance of country *i* from South Africa. $X_{4,it},\cdots,X_{k,it}$ represent other factors from the economic growth literature.

Independent variables	Coefficient	(i)	(ii)	(iii)
SAF	β_1	0.55^{***}	0.62^{***}	0.53^{***}
$(SAF)(TRADE)$	β_2		0.21^{***}	0.20^{***}
$(SAF)(DIST)$	β_3			-0.15
Other factors		included	included	included

*** and ** indicate the coefficient is significant at the level of 1% and 5%, respectively.

(a) To make a conclusion about whether the South African economic growth influenced the other African countries, which regression model is most appropriate?
(Use a 5% significance level for testing hypotheses.)

(b) Based on the regression chosen in (a), what is the net effect of the South African economic growth (SAF) on the other countries ($\%\Delta GDP_{it}$)?

(c) Is the net effect constant?

Problem 4 (empirical)

Some U.S. states have enacted laws that allow citizens to carry concealed weapons. These laws are know as "shall- issue" laws because they instruct local authorities to issue a concealed weapons permit to all applicants who are citizens, are mentally competent, and have not been convicted of a felony. (Some states have additional restrictions.) Proponents argue that if more

people carry concealed weapons, crime will decline because criminals will be deterred from attacking other people. Opponents argue that crime will increase because of accidental or spontaneous use of the weapons. In this exercise you will analyze the effect of concealed weapons on violent crimes. Use the data set Guns.csv (on Latte) which contains a balanced panel of data from the 50 U.S. states plus the District of Columbia for 1977-1993. **A detailed description of all variables included is at the end of this Problem set.**

a. Estimate (1) a regression of ln(vio) against shall and (2) a regression of ln(vio) against shall, incarc_rate, density, avginc, pop, pb1064, pw1064, and pm1029. Interpret the coefficient on shall in regression (2). Is this estimate economically significant?

b. Does adding the control variables in regression (2) change the estimated effect of a shall-carry law in regression (1)? Is the coefficient statistically significant?

c. Suggest a variable that varies across states but plausibly varies little, or not at all, over time and that could cause omitted variable bias in regression (2). Explain how this variable fulfills both conditions for OVB.

d. Do the results change when you add state fixed effects? If so, which set of regression results is more credible, and why?

e. Do the results change when you add time fixed effects (in addition to the state fixed effects)? If so, which set of results is more credible, and why?

f. In your view, what is the most important remaining threat to the internal validity of this regression analysis? Explain briefly.

g. Based on your analysis, what conclusions would you draw about the effects of concealed weapons laws on the violent crime rate?

Variable Definitions

Variable	Definition
vio	violent crime rate (incidents per 100,000 members of the population)
rob	robbery rate (incidents per 100,000)
mur	murder rate (incidents per 100,000)
shall	= 1 if the state has a shall-carry law in effect in that year = 0 otherwise
incarc_rate	incarceration rate in the state in the previous year (sentenced prisoners per 100,000 residents; value for the previous year)
density	population per square mile of land area, divided by 1000
avginc	real per capita personal income in the state, in thousands of dollars
pop	state population, in millions of people
pm1029	percent of state population that is male, ages 10 to 29
pw1064	percent of state population that is white, ages 10 to 64
pb1064	percent of state population that is black, ages 10 to 64
stateid	ID number of states (Alabama = 1, Alaska = 2, etc.)
year	Year (1977-1999)