Problem set 5 solutions

1. Answer:

   (a) TRUE. The statement is about equivalence of the estimators, so it does not matter if the exogeneity assumption is violated. Take the average of both sides of the fixed effects regression model we get $\bar{Y}_i = \beta_1 \bar{X}_i + \alpha_i + \bar{u}_i$. The difference between the fixed effects regression and the average regression implies a regression model $\tilde{Y}_{it} = \beta_1 \tilde{X}_{it} + \tilde{u}_{it}$, which is the same as the demeaned regression model. Thus, their respective OLS estimators should agree.

   Alternatively, consider the first-order condition for $\hat{\beta}_1$:

$$\sum_{i=1}^{n}\sum_{t=1}^{T} \hat{\beta}_1 (Y_{it} - \hat{\alpha}_i - \hat{\beta}_1 X_{it}) = 0 \Rightarrow$$
$$\sum_{i=1}^{n} \hat{\beta}_1 (\frac{1}{T}\sum_{t=1}^{T} Y_{it} - \hat{\alpha}_i - \hat{\beta}_1 \frac{1}{T}\sum_{t=1}^{T} X_{it}) = 0 \Rightarrow$$
$$\sum_{i=1}^{n} \hat{\beta}_1 (\tilde{Y}_{it} - \hat{\beta}_1 \tilde{X}_{it}) = 0$$

   where the last line is the difference between the first two lines. The last line is the same first-order condition for $\hat{\beta}_2$. Thus, the two estimators are equivalent.

   (b) FALSE. The interactions of every entity dummy with every time dummy would perfectly explain the data. Every data point will give us $\hat{\alpha}_{it} = Y_{it}$ with no residual. So, there would be no residual variation to explain with X_{it}.

   (c) FALSE. We would have perfect multi-collinearity if we include a constant variable and one dummy variable for each entity. Therefore, we only need to exclude the dummy variable for one of the entities when we include an intercept. If we do not include the intercept, then we can include all dummy variables.

   (d) FALSE. LAT is constant across time for a country (assuming border changes over the past 20 years do not affect average latitude of countries affected). We do not need to control for LAT as LAT is perfectly collinear with the entity fixed effects.

2. Answer:

   (a) Regression (2) is preferable as there may be autocorrelation in the erros terms for each state. Clustered standard errors are also robust to heteroskedasticity.

   (b) A one percentage point increase in the unemployment rate in a recession year decreases the vote share of the incumbent by 0.66 percentage points. In a non-recession year, a one percentage point increase in the unemployment rate increases the vote share by 1.32 percentage points.

   (c) We can see that the unemployment rate does affect incumbent vote share both from the fact that UR is statistically significant in regression (3) and from the fact that the F-statistic for the joint significance of UR and UR×Non-rec year is 4.58.

   (d) Regression (5) captures non-linearities in the effects of the unemployment rate by adding a quadratic term on UR. Given that the coefficients on the non-linear terms cannot be statistically distinguished from zero, regression (4) is preferred to avoid overfitting.

(e) An example of a variable not explicitly in regression (4) which could cause omitted variable bias but which is controlled for by including time fixed effects is whether an incumbent president cut taxes during their term. Tax cuts are politically popular and will raise the incumbent vote share. They may also reduce unemployment by generally boosting economic activity. However, given that tax cuts are federal, they would affect every state in the same way and be captured by the time fixed effects.

(f) An example of a variable not explicitly in regression (4) which could cause omitted variable bias but which is controlled for by including state fixed effects is the share of a state that is urban vs. rural. If rural areas are less targetted by federal policies, for example, or if they consume less mass media, then rural areas would be less susceptible to an incumbent advantage. Rural areas tend to also have higher unemployment than urban areas. In general, how urban vs. rural a state is, is fixed over time. Adding state fixed effects would control for this potentially omitted variable.

(g) The effects of local economic conditions (state unemployment rates) on incumbent vote shares is dependent on whether or not the U.S. economy is in a recession. During recessions, states with higher unemployment rates see less votes given to the incumbent. This effect is reversed when there is no recession.

3. Answer:

(a) In 1982 we see that a 1 percentage point increase in unemployment is associated with 1.3 more crimes per 1000 people. In 1987, we see that the same increase is associated with 4.2 fewer crimes per 1000 people. The 1982 result matches intuition, but the 1987 result is unexpected.

(b) The natural resource capacity of a city (i.e. forests, gas, oil, etc) affects the wealth of a city. Cities with more plentiful natural resources may have lower unemployment and lower crime due to this increased wealth and productive capacity, which could cause omitted variable bias in (a).

(c) The results change such that we observe a null effect. The point estimate is -0.018 crimes per 1000 people, economically insignificant, with a large standard error of 0.628. The regression is

$$\text{CrimeRate}_{it} = \alpha_i + \beta \text{Unemp}_{it} + \varepsilon_{it}$$

(d) The generosity of national welfare programs may be such an omitted variable. For example, a reduction in federal welfare may reduce unemployment by forcing workers to accept lower wages to survive. At the same time, the reduction in welfare may also lead to more crime as people have fewer resources. This would cause downwards bias in a regression without time fixed effects, but would be fully accounted for by time fixed effects as federal welfare program rules are the same for all cities.

(e) When we include both time and city fixed effects we find that a 1 percentage point higher unemployment rate is associated with 2.22 more crimes per 1000 people (t-stat = 2.72). This is the most credible result, as it is comparing crime rates between cities, holding fixed the permanent properties of a city and the time-dependent factors that affect all cities. The regression specification is

$$\text{CrimeRate}_{it} = \alpha_i + \gamma_t + \beta \text{Unemp}_{it} + \varepsilon_{it}$$

(f) I would conclude that that unemployment and crime are positively related, based on the results of the second fixed effects regression.