EC3303 Econometrics I Tutorial Problem Set 3

1. A researcher plans to study the causal effect of a strong legal system on the economy, using data from a sample of countries. The researcher plans to regress national income per capita on whether the country has a strong legal system or not (an indicator variable, based on expert opinion taking the value 1 if the country has a strong legal system and taking the value 0 otherwise).
   Do you think this regression suffers from omitted variable bias? Why or why not? Which variables would you add to the regression if you think the regression suffers from omitted variable bias?

   Answer:

   Yes, the OLS estimator of the effect of a strong legal system on per capita national income is likely to be biased. This is because we have omitted from the regression, variables which influence national income and which are potentially also correlated with the strength of the legal system.
   Omitted variables could include things such as the level of the capital stock, the level of technological development, etc (creativity is encouraged here).

2. A researcher collects data from a random sample of 420 dwellings in the Bukit Timah area in 2015. Let *Price* denote the selling price (in thousands of Singapore dollars), *BDR* denote the number of bedrooms, *Bath* denote the number of bathrooms, *Hsize* denote the size of the house (in square meters), *Age* denote the age of the house (in years), and *Poor* denote a binary variable that is equal to 1 if the condition of the house is reported as "poor". An estimated regression yields (with robust standard errors in parentheses):

$$\underset{}{\widehat{Price}} = \underset{}{110.7} + \underset{(22.5)}{0.425}BDR + \underset{(2.43)}{22.5}Bath + \underset{(8.45)}{1.69}Hsize + \underset{(0.112)}{0.095}Age - \underset{(0.322)}{47.2}Poor, \quad (10.7)$$

$$\bar{R}^2 = 0.692, SER = 42.5$$

(a) Suppose that a homeowner in Bukit Timah adds a new bathroom to her house, which increases the size of the house by 8.7 square meters. What is the predicted increase in the value of the house, other things being equal?

Answer: $\Delta Bath = 1$ and $\Delta Hsize = 8.7$. The expected change in price is 22.5+1.69(8.7)=37.2 thousand dollars or $37,200.

(b) What is the predicted loss in the value of the house if a homeowner lets her house run down, such that its condition becomes "poor", other things being equal?

Answer: The loss is $47,200.

(c) Interpret the coefficient on BDR. Are the results from this regression consistent with the claim that houses with more bedrooms sell for a higher price?

Answer: The coefficient on *BDR* measures the effect on price of a unit increase in the number of bedrooms, holding *Bath*, *Hsize*, *Age*, and *Poor* constant.

The coefficient on *BDR* is not statistically significantly different from zero at the 5% level (t-statistic=0.425/2.43=0.175<1.96). Hence, the results are inconsistent with the claim that houses with more bedrooms sell for a higher price.

(d) Construct a 99% confidence interval for the effect of a one year increase in the age of a house. Round your answer to 3 decimal places.

Answer:
99% CI = $0.095 \pm (2.58 \times 0.322) = (-0.736, 0.926)$.

3. The cost of attending your university has once again gone up. Although you have been told that education is investment in human capital, which carries a return of roughly 10% a year, you are not pleased. One of the administrators at your university does not make the situation better by telling you that you pay more because the reputation of your institution is better than that of others. To investigate this hypothesis, you randomly collect data on 100 U.S. universities and liberal arts colleges from the 2000-2001 *United States News and World Report* annual rankings. Next, you perform the regression

$$\widehat{Cost} = \underset{(2,058.63)}{7,311.17} + \underset{(664.58)}{3,985.20} \times Reputation - \underset{(0.13)}{0.27} \times Size$$

$$+ \underset{(2,154.85)}{8,406.79} \times Dpriv - \underset{(1,121.92)}{416.38} \times Dlibart - \underset{(1,007.86)}{2,376.51} \times Dreligion$$

$$R^2 = 0.72, \; SER = 3,773.35$$

where *Cost* is tuition fees per year in U.S. dollars; *Reputation* is the index used in the

U.S. News and World Report (based on a survey of university presidents and chief academic officers), which ranges from 1 ("poor") to 5 ("excellent"); *Size* is the number of undergraduate students; and *Dpriv*, *Dlibart*, and *Dreligion* are binary variables indicating whether the institution is private, a liberal arts college, and has a religious affiliation respectively.

(a) What is the forecasted cost for a liberal arts college, which has no religious affiliation, a size of 1,500 students and a reputation level of 4.5? Note that all liberal arts colleges are private (but not all private schools are liberal arts colleges).

Answer: $32,830.

(b) Your friend, Joe is studying in the United States. To save money, he is willing to switch from a private university to a public university, which has a reputation ranking of 0.5 less and 10,000 more students. How do you predict his cost will change?

Answer: His cost is predicted to decrease by $13,099 per year.

(c) You want to test simultaneously the hypotheses that $\beta_{size} = 0$ and $\beta_{Dilbart} = 0$. STATA returns an *F*-statistic of 5.23. What do you conclude?

Answer: The critical value for $F_{2,\infty}$ is 3.00 (5% level) and 4.61 (1% level). Since F-statistic = 5.23 $>$3.00, we reject the null hypothesis that $\beta_{size} = 0$ and $\beta_{Dilbart} = 0$ at the 5% level.

Stata Exercise (to be done in tutorial with the tutor)

4. The California Standardized Testing and Reporting data set (caschool.dta) contains data on test performance, school characteristics, and student demographic backgrounds for 420 Californian school districts. Today, we will learn how to perform linear regression with multiple regressors and demonstrate how perfect multicollinearity could arise.