BUEC 333-D100, Test 2

July 27, 2016

1 Chapter 6

Using data on the SAT scores, personal and high school characteristic of 4000 U.S. students, we estimate the following regression:

$$\widehat{SAT} = 1028 - 2.19Hsize - 45.09Female$$
$$(2.06) \qquad (3.37)$$

where SAT is the student's SAT score, $Hsize$ is the number of students in their high school graduating class (in hundreds), $Female$ is a gender dummy (1 if student is female, 0 otherwise).

1. Interpret the coefficient estimate of -2.19 on $Hsize$. **If Hsize goes up by one unit, then we estimate that the expected SAT score goes down by 2.19, ceteris paribus. Larger high school cohorts are bad for you SAT score?**

2. Is there sufficient statistical evidence that the size of your high school graduating class has an effect on your SAT score? **No, |2.19/2.06|<1.96**

3. Interpret the coefficient estimate of -45.09 on $Female$. **We estimate that, ceteris paribus, women score 45.09 points lower, in expectation.**

Next, we control for whether the student is part of a visible minority. Let $VisMin$ is a dummy variable that is 1 if the student is part of a visible minority, and 0 otherwise.

$$\widehat{SAT} = 748 - 1.68Hsize - 14.09Female - 69.17VisMin$$
$$(2.13) \qquad (4.12) \qquad (10.87)$$

where

4. Is there any evidence of omitted variable bias? Explain! **Yes: the coefficient on female changes drastically when we include VisMin.**

5. What can you say about the relationship between $Female$ and $Vismin$ based on the two regression results? **In our sample with SAT result, women are more likely to be visible minorities.**

6. Interpret your result in 5. **An example of a correct answer: the proportion of SAT test takers that are women is higher among visible minorities than for non-(visible minorities).**

2 Chapter 5+8

Suppose that we have data on BUEC 333 test scores (Y_i), duration for which student i studies for exam (X_i), and the major of the student, call it D_i, where

$$D_i = \begin{cases} 1, \text{ if economics major} \\ 0, \text{ if non economics major} \end{cases}$$

Consider the following model:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 D_i + \beta_3 D_i X_i + u_i \tag{1}$$

where Assumption 1 holds:

$$E\left(u_i | X_i, D_i\right) = 0. \tag{2}$$

Y_i is the score between 0 and 100. X_i is the duration studied in hours, between 1 and 100.

1. What is the intercept for econ-majors? **Set Di=1 and Xi=0 to find that the intercept is beta0+beta2.**
2. What is the slope for econ-majors? **beta1+beta3**
3. What is the interpretation (in words) of β_1? **Conditional on not being an econ major, and keeping other things constant, beta1 is the increase in the conditional expectation of your test score for an additional hour of study.**
4. What is the interpretation (in words) for β_3? **Additional - on top of beta1 - c.p. increase in the conditional expectation of your test score if you are an econ major.**

Now suppose that the OLS estimate for β_1 is 0.60, and its standard error is 0.2.

5. Test whether $\beta_1 = 0$, at the 5% and 1% significance level. **Reject in both cases, since |0.6/0.2|>2.58>1.96.**

Finally, we try a different model:

$$\ln(Y_i) = \beta_0 + \beta_1 \ln(X_i) + \beta_2 D_i + \beta_3 D_i \ln(X_i) + u_i \tag{3}$$

6. What is the interpretation of β_1 in this model? **For non-econ majors, the expected percentage increase in their score for a 1 percent increase in their hours studied, c.p.**

3 Chapter 9

Solution: see chapter 9 of the textbook.

1. What is the difference between internal and external validity?
2. List the five threats to internal validity.
3. Explain in your own words one of the five threats mentioned under (2).

4 Chapter 12

Solution: the solutions to this question can be found in Sections 12.4 and 12.5 of the textbook!

Your textbook's Section 12.4 and 12.5 mentions a few situations which require IV techniques. Some of those are:

- the effect of class size on test scores
- the effect of an increase in the prison population on crime rates
- the aggressive treatment of heart attacks and the potential for saving lives
- the demand elasticity for cigarettes

1. Pick one of these examples and explain what each variable in the model below represents:

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

That is, in the example you choose, what is Y_i and X_i?

2. The solution to the issue in (1) is to find an instrumental variable Z_i. What is an example of an IV that the book mentions?
3. What are the two conditions that this instrumental variable must satisfy?