

Recitation 4

Sean Hyland

UN3412 Introduction to Econometrics

Spring 2020

Outline

1. Q1: S&W Additional Empirical Exercise 5.2
2. Q2: S&W Additional Empirical Exercise 5.3
3. Q3: S&W Additional Empirical Exercise 6.1
4. Q4: S&W Additional Empirical Exercise 6.2

Outline

1. Q1: S&W Additional Empirical Exercise 5.2
2. Q2: S&W Additional Empirical Exercise 5.3
3. Q3: S&W Additional Empirical Exercise 6.1
4. Q4: S&W Additional Empirical Exercise 6.2

Q1: S&W Additional Empirical Exercise 5.2 (i)

Using the data set `TeachingRatings` described in Empirical Exercise AEE4.2, run a regression of `Course_Eval` on `Beauty`.

- (a) Is the estimated regression slope coefficient statistically significant? That is, can you reject the null hypothesis $H_0 : \beta_1 = 0$ versus a two-sided alternative at the 10%, 5%, or 1% significance level?
- (b) What is the p-value associated with coefficient's t-statistic?

Outline

1. Q1: S&W Additional Empirical Exercise 5.2
2. Q2: S&W Additional Empirical Exercise 5.3
3. Q3: S&W Additional Empirical Exercise 6.1
4. Q4: S&W Additional Empirical Exercise 6.2

Q2: S&W Additional Empirical Exercise 5.3 (i)

Using the data set *CollegeDistance* described in Empirical Exercise AEE4.3, run a regression of years of completed education (*ED*) on distance to the nearest college (*Dist*) and carry out the following exercises.

- (a) Is the estimated regression slope coefficient statistically significant? That is, can you reject the null hypothesis $H_0 : \beta_1 = 0$ versus a two-sided alternative at the 10%, 5%, or 1% significance level? What is the p-value associated with coefficient's t-statistic?
- (b) Construct a 95% confidence interval for the slope coefficient.
- (c) Run the regression using data only on females and repeat (b).
- (d) Run the regression using data only on males and repeat (b).
- (e) Is the effect of distance on completed years of education different for men than for women? (Hint: See Exercise 5.15.)

Outline

1. Q1: S&W Additional Empirical Exercise 5.2
2. Q2: S&W Additional Empirical Exercise 5.3
3. Q3: S&W Additional Empirical Exercise 6.1
4. Q4: S&W Additional Empirical Exercise 6.2

Q3: S&W Additional Empirical Exercise 6.1 (i)

Using the data set *TeachingRatings* described in Empirical Exercise AEE4.2, carry out the following exercises.

- (a) Run a regression of *Course_Eval* on *Beauty*. What is the estimated slope?
- (b) Run a regression of *Course_Eval* on *Beauty*, including some additional variables to control for the type of course and professor characteristics. In particular, include as additional regressors *Intro*, *OneCredit*, *Female*, *Minority*, and *NNEnglish*.¹ What is the estimated effect of *Beauty* on *Course_Eval*? Does the regression in (a) suffer from important omitted variable bias?
- (d) Professor Smith is a black male with average beauty and is a native English speaker. He teaches a three-credit upper-division course. Predict Professor Smith's course evaluation.

¹See https://wps.pearsoned.com/wps/media/objects/11422/11696965/aee/TeachingRatings_Description.pdf for a description of each variable.

Outline

1. Q1: S&W Additional Empirical Exercise 5.2
2. Q2: S&W Additional Empirical Exercise 5.3
3. Q3: S&W Additional Empirical Exercise 6.1
4. Q4: S&W Additional Empirical Exercise 6.2

Q4: S&W Additional Empirical Exercise 6.2 (i)

Using the data set *CollegeDistance* described in Empirical Exercise AEE4.3, carry out the following exercises.

- (a) Run a regression of years of completed education (ED) on distance to the nearest college ($Dist$). What is the estimated slope?
- (b) Run a regression of ED on $Dist$, but include some additional regressors to control for characteristics of the student, the student's family, and the local labor market. In particular, include as additional regressors *Bytest*, *Female*, *Black*, *Hispanic*, *Incomehi*, *Ownhome*, *DadColl*, *Cue80*, and *Stwmfg80*.² What is the estimated effect of $Dist$ on ED ?
- (c) Is the estimated effect of $Dist$ on ED in the regression in (b) substantively different from the regression in (a)? Based on this, does the regression in (a) seem to suffer from important omitted variable bias?
- (d) Compare the fit of the regression in (a) and (b) using the regression standard errors, R^2 and \bar{R}^2 . Why are the R^2 and \bar{R}^2 so similar in regression (b)?
- (e) The value of the coefficient on *DadColl* is positive. What does this coefficient measure?

Q4: S&W Additional Empirical Exercise 6.2 (ii)

- (f) Explain why *Cue80* and *Stwmfg80* appear in the regression. Are the signs of their estimated coefficients (+ or -) what you would have believed? Interpret the magnitudes of these coefficients.
- (g) Bob is a black male. His high school was 20 miles from the nearest college. His base-year composite test score was 58. His family income in 1980 was \$26,000, and his family owned a home. His mother attended college, but his father did not. The unemployment rate in his county was 7.5%, and the state average manufacturing hourly wage was \$9.75. Predict Bob's years of completed schooling using the regression in (b).
- (h) Jim has the same characteristics as Bob except that his high school was 40 miles from the nearest college. Predict Jim's years of completed schooling using the regression in (b).

²See https://wps.pearsoned.com/wps/media/objects/11422/11696965/aee/CollegeDistance_DataDescription.pdf for a description of each variable.