Econometrics

Question 9 (Econometrics, 60 points).

This question consists of two subquestions, each worth 30 points.

1. **(30 points)** In this question we will look at the relation between the logarithm of weekly earnings and years of education. Using data from the National longitudinal study of youth, we find the following results for a regression of log weekly wage on years of education, experience, experience squared and an intercept:

$$\begin{array}{lccccccc} \widehat{\log(\text{earnings})}_i & = & 4.016 & + & 0.092 \cdot \text{educ}_i & + & 0.079 \cdot \text{exper}_i & - & 0.002 \cdot \text{exper}_i^2 \\ & & (0.222) & & (0.008) & & (0.025) & & (0.001) \end{array}$$

a. **(5 points)** Construct a 95% confidence interval for the effect of years of education on log weekly earnings.

Answer:

$[0.092 - 1.96 \times 0.008, 0.092 + 1.96 \times 0.008]$

Points:

1 point for noting the mean is 0.092.

1 point for noting the standard deviation is 0.008.

1 point for using 1.96.

2 points for a correct answer. (Students do not have to calculate the final numerical result $[0.07632, 0.10768]$ to be given full credit.)

b. **(6 points)** Consider an individual with 10 years of experience. What would you expect to be the return to an additional year of experience for such an individual (the effect on log weekly earnings)?

Answer:

Experience changes from 10 to 11, experience squared changes from 100 to 121, so total effect is $0.079 - 21 \times 0.002 = 0.037$

Points:

1 point for noting exper changes from 10 to 11.

2 points for noting exper^2 changes from 100 to 121.

2 points for a correct equation $0.079 - 21 \times 0.002$.

1 point for a correct answer 0.037.

c. **(4 points)** Labor economists studying the relation between education and earnings are often concerned about what they call "ability bias." Suppose that invididuals differ in ability, and that the correct specification of the regression function is one that includes ability:

$$\log(\text{earnings})_i = \beta_0 + \beta_1 \cdot \text{educ}_i + \beta_2 \cdot \text{exper}_i + \beta_3 \cdot \text{exper}_i^2 + \beta_4 \cdot \text{ability}_i + \varepsilon_i$$

In this regression, what do you expect the sign of β_4 (the coefficient on ability) to be?

Answer:

Expect this to be positive.

Points:

4 points for a correct answer.

Partial credit for explanation (but wrong answer) at the grader's own judgement.

d. **(4 points)** What do you think is the sign of the correlation between ability and years of education?

Answer:

Positive correlation between ability and years of education.

Points:

4 points for a correct answer.

Partial credit for explanation (but wrong answer) at the grader's own judgement.

e. **(6 points)** If we estimate the regression function with ability included, do you think that the estimated value of β_1 will be greater or less than what it was in the regression without ability? Explain.

Answer:

Omitted variable bias is $\beta_4 \cdot \text{cov}(\text{ability}, \text{educ})/\text{var}(\text{educ})$, which is positive. Therefore the estimated value of β_1 will be less than in the previous regression.

Points:

2 points for a correct answer (estimated value of β_1 is smaller).

4 points for a good explanation.

Partial credit at the grader's own judgement.

f. **(5 points)** Another approach to dealing with ability bias has been to use data on pairs of individuals from the same household, or even better, twins. Suppose you had observations for N households with twins. Let Y_{i1} and Y_{i2} be the values of log earnings for the first and second twin in household i, and X_{i1} and X_{i2} be their education level. What regression would you run to estimate the effect of education on earnings to avoid ability bias?

Answer:

Regress $(Y_{i2} - Y_{i1})$ on $(X_{i2} - X_{i1})$.

Points:

5 points for a correct answer.

Partial credit at the grader's own judgement.

2. **(30 points)** To estimate the effect of minimum wage on employment, David Card and Alan Krueger, collected data on fast food restaurants in two neighbouring states, New Jersey and Pennsylvania, in 1992, before and after an increase in the legal minimum wage came into effect in New Jersey. Prior to that the minimum wage was the same in New Jersey and Pennsylvania. Let Y_i be employment in restaurant i, let S_i be an indicator for the restaurant being in New Jersey, and let T_i be an indicator for whether the employment was measured post minimum wage change.

   a. **(5 points)** Suppose we only had data from New Jersey. What is the regression you would run to estimate the effect of the change in the minimum wage?

   Answer:

   Regress Y_i on a constant and a dummy for the time period. Leads to average employment in NJ post minus average employment in NJ prior to change.

   Points:

   1 point for noting the dependent variable is Y_i.

   1 point for noting the constant.

   2 points for noting the other regressor is time dummy.

   1 point for a correct answer (regress Y_i on constant and time dummy). The second sentence is not a requisite for full credit.

   Partial credit at the grader's own judgement.

   b. **(5 points)** Suppose we only had data on employment after the change in the minimum wage, in both states. What regression would you run?

   Answer:

   Regress Y_i on constant and S_i. Leads to average post-change employment in NJ minus average post-change employment in Pennsylvania.

   Points:

   1 point for noting the dependent variable is Y_i.

1 point for noting the constant.

2 points for noting the other regressor is S_i.

1 point for a correct answer (regress Y_i on constant and S_i). The second sentence is not a requisite for full credit.

Partial credit at the grader's own judgement.

c. **(6 points)** What are the problems with both these estimation strategies?

Answer:

Permanent differences between NJ and PA for (b), and other changes over time in (a).

Points:

3 points for correctly pointing out the problem for (a).

3 points for correctly pointing out the problem for (b).

Partial credit at the grader's own judgement.

d. **(7 points)** What regression would you run with both before and after data for both New Jersey and Pennsylvania?

Answer:

Regress Y_i on constant, time indicator, state indicator, and interaction between time and state.

Points:

1 point for noting the dependent variable is Y_i.

1 point for noting the constant.

1 point for noting the time indicator.

1 point for noting the state indicator.

2 points for noting the interaction between time and state.

1 point for a correct answer (regress Y_i on constant, time indicator, state indicator, and interaction between time and state).

Partial credit at the grader's own judgement.

e. **(7 points)** Suppose we had data from three states, the third one (New York) like Pennsylvania not experiencing a change in the legal minimum wage. What regression would you do now?

Answer:

Regress Y_i on a constant, dummies for PA, NY, and time dummy, and interaction of NJ and time dummy.

Points:

1 point for noting the dependent variable is Y_i.

1 point for noting the constant.

2 points for noting the dummies for PA and NY.

1 point for noting the time dummy.

1 point for noting the interaction between NJ and time dummy.

1 point for a correct answer (regress Y_i on a constant, dummies for PA, NY, and time dummy, and interaction of NJ and time dummy).

Partial credit at the grader's own judgement.