EC 421

Final

Spring 2019

Full Name ←

UO ID ←

Total points: /108

No phones, calculators, or outside materials.

True/False

32.5 points

Note: You do not need to explain to your answers **in this section**.

01. [T/F] (**2.5pts**) In the presence of omitted-variable bias, ordinary least squares (OLS) is still consistent.

02. [T/F] (**2.5pts**) In the model $\text{Births}_t = \beta_0 + \beta_1 \text{Income}_t + u_t$, only the current period's income affects the current period's number of births.

03. [T/F] (**2.5pts**) In the model $\text{Births}_t = \beta_0 + \beta_1 \text{Income}_t + \beta_2 \text{Income}_{t-1} + u_t$, only the current period's income affects the current period's number of births.

04. [T/F] (**2.5pts**) In the model $\text{Births}_t = \beta_0 + \beta_1 \text{Income}_t + \beta_2 \text{Births}_{t-1} + u_t$, only the current period's income affects the current period's number of births.

05. [T/F] (**2.5pts**) If an estimator is biased, then it is not consistent.

06. [T/F] (**2.5pts**) As long as the mean of a variable does not change with time, the variable is stationary.

07. [T/F] (**2.5pts**) With a *p*-value of 0.049 and a significance level of 0.05, we reject the null hypothesis.

08. [T/F] (**2.5pts**) Random walk walks are stationary.

09. [T/F] (**2.5pts**) If the average *untreated* outcome ($textY_{0i}$) in the treated group equals the average *untreated* outcome in the control group, then we do not need to worry about selection bias.

10. [T/F] (**2.5pts**) Exogeneity requires $E\left[u_i^2\right] = \sigma^2$.

11. [T/F] (**2.5pts**) Randomized experiments typically have issues with selection bias.

12. [T/F] (**2.5pts**) Causation has nothing to do with correlation.

13. [T/F] (**2.5pts**) Instrumental variables estimates can be consistent when OLS is inconsistent.

Short definitions

33 points

14. (**3pts**) Define the "standard error of an estimator".

15. (**3pts**) Define "autocorrelation".

16. (**3pts**) Define "nonstationarity".

17. (**3pts**) Define "p-value".

18. (**3pts**) Define "causality".

19. (**3pts**) Define "selection bias".

20. (**3pts**) Define "the fundamental problem of causal inference"

21. (**3pts**) What is a "random walk"?

22. (**6pts**) What are the two requirements for a *valid* instrument? Briefly define each requirement.

23. (**3pts**) What is "heteroskedasticity"?

Short answer

42.5 points

24. Suppose you want to estimate the causal effect of education on future earnings.

a. (**6pts**) Explain why regressing earnings on education is not going to give you the answer you want (*i.e.*, the causal effect of education on earnings).

b. (**4pts**) Now imagine you know there is a scholarship program that randomly gives out scholarships. Individuals who receive these scholarships. Explain how you could use instrumental variables to estimate the causal effect of education on earnings.

☐ Check this box if you want to skip all of question 25 and instead receive **2pts**.

25. Each part of this question refers to the following `R` output which results from estimating

```
lm(income ~ education + experience + female, data = wage_df)
```

```
#> # A tibble: 4 x 5
#>   term        estimate std.error statistic  p.value
#>   <chr>          <dbl>     <dbl>     <dbl>    <dbl>
#> 1 (Intercept)    3177.    2606.       1.22 2.26e- 1
#> 2 education       952.     109.       8.74 7.66e-14
#> 3 experience     2898.      89.0     32.6  1.10e-53
#> 4 female        -1761.    1570.      -1.12 2.65e- 1
```

a. (**2pts**) Write down the model that we've estimated (do not include the estimates, just the βs).

b. (**3pts**) Carefully interpret the coefficient on `education`.

c. (**3pts**) Conduct a hypothesis test for the coefficient on `education`. Describe each step and the explain your conclusion.

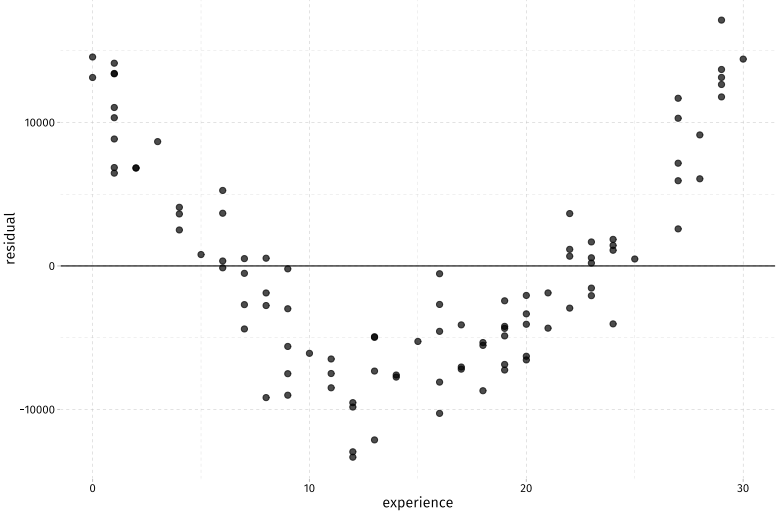
d. (**3pts**) Carefully interpret the coefficient on `female` (an indicator variable for *female*).

e. (**2.5pts**) **True/False** Because the p-value on `female` is greater than 0.05, we can conclude there is no gender-based difference in income.

f. (**2pts**) If we remove `female` from the regression, do you expect the R^2 to increase or decrease? Explain your answer.

g. (**3pts**) What assumptions/conditions must be satisfied for us to be able to interpret the coefficient on `education` as causal?

h. (**2pts**) Examine the following plot of the residuals from the previous regression (on the y axis) and `experience` (on the x axis). Describe any issues that suggests and how we might 'fix' them.



☐ Check this box if you want to skip all of question 26 and instead receive **2pts**.

26. Consider the following model for the number of alcoholic drinks an individual consumes in a day.

$$\mathbf{Drinks}_t = \beta_0 + \beta_1 \mathbf{Drinks}_{t-1} + \beta_2 \mathbf{Income}_t + \beta_3 \mathbf{Income}_{t-1} + u_t$$

a. (**2pts**) Carefully interpret the term β_1.

b. (**2pts**) Carefully interpret the term β_2.

c. (**2pts**) Carefully interpret the term β_3.

d. (**2pts**) What does $\beta_2 + \beta_3$ tell us?

e. (**2pts**) What does this model assume about the effect of income throughout time?

f. (**2pts**) What happens to our OLS-based estimate of β_1 if u_t is autocorrelated?