EC 421

Final

18 March 2019

Full Name ←

UO ID ←

No phones, calculators, or outside materials.

True/False

37.5 points

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

01. [T/F] (**2.5pts**) $\text{Births}_t = \beta_0 + \beta_1 \text{Income}_t + \beta_2 \text{Births}_{t-1} + u_t$ describes a static time series model.

02. [T/F] (**2.5pts**) In the model $\text{Births}_t = \beta_0 + \beta_1 \text{Income}_t + u_t$, the parameter β_1 gives the effect of income in **all** previous time periods on births in the current period.

03. [T/F] (**2.5pts**) From the estimated model

$$\text{Births}_t = \hat{\beta}_0 + \hat{\beta}_1 \text{Income}_t + \hat{\beta}_2 \text{Income}_{t-1} + \hat{\beta}_3 \text{Income}_{t-2} + e_t$$

we can estimate the *total* effect of income on births as $\hat{\beta}_0 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3$.

04. [T/F] (**2.5pts**) If the disturbances from two time periods t and s (with $t \neq s$) have non-zero covariance, *i.e.*, $\text{Cov}(u_t,\, u_s) \neq 0$, then we say that the disturbance is *heteroskedastic*.

05. [T/F] (**2.5pts**) For dynamic models with lagged explanatory variables and autocorrelated disturbances, OLS we can trust OLS to be unbiased.

06. [T/F] (**2.5pts**) If $\text{Var}(u_t) = 1$ for t in {1, ..., 10} and $\text{Var}(u_t) = 3$ for {11, ..., T}, then u_t is heteroskedastic.

07. [T/F] (**2.5pts**) If $\text{Var}(u_t) = 1$ for t in {1, ..., 10} and $\text{Var}(u_t) = 3$ for {11, ..., T}, then u_t is nonstationary.

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

09. [T/F] (**2.5pts**) Selection bias refers to including observations with differing variances.

10. [T/F] (**2.5pts**) *Prediction* focuses on estimating $\hat{\beta}$, while *casual inference* focuses on estimating $\hat{y}$.

11. [T/F] (**2.5pts**) In the Rubin causal model, $y_{1,i}$ refers to the outcome for individual i when he/she does not receive treatment.

12. [T/F] (**2.5pts**) Randomized experiment help us avoid selection bias by (approximately) balancing $Avg(y_{1,i}|D_i = 1)$ and $Avg(y_{1,i}|D_i = 0)$.

13. [T/F] (**2.5pts**) An instrumental variable is *exogenous* if it affects y (the outcome) through the endogenous x and through another explanatory variable w.

14. [T/F] (**2.5pts**) The relevance requirement of instrumental variables is untestable.

15. [T/F] (**2.5pts**) The exogeneity requirement of instrumental variables is untestable.

Short Answer

62.5 points

Note: You will typically need to explain/justify your answers in this section.

16. (**2.5pts**) Write down an ADL(1,1) model for the effect of income on births.

17. (**2.5pts**) Explain what negative autocorrelation in our disturbances means.
You can focus on an AR(1) process.

18. (**3pts**) In the following dynamic time-series model, u_t is first-order autocorrelated, *i.e.*,

$$\begin{aligned} \text{Health}_t &= \beta_0 + \beta_1 \text{Income}_t + \beta_2 \text{Health}_{t-1} + u_t && \text{model, } t \\ \text{Health}_{t-1} &= \beta_0 + \beta_1 \text{Income}_{t-1} + \beta_2 \text{Health}_{t-2} + u_{t-1} && \text{model, } t-1 \\ u_t &= \rho u_{t-1} + \varepsilon_t && \text{AR(1)} \end{aligned}$$

where ε_t is "white noise"—independentaly and identically distributed with mean zero and variance σ_ε^2.

Explain why OLS will likely be biased for β_2—even if there are no omitted variables.

19. (**3pts**) Suppose we are concerned about first-degree autocorrelation for the model

$$\text{Crime}_t = \beta_0 + \beta_1 \text{Police}_t + \text{Crime}_{t-1} + u_t$$

To test whether u_t has first-degree autocorrelation, we run a Breusch-Pagan test and find an F statistic of 3.98, which has a p-value of approximately 0.0488. State the conclusion of this test and interpret the results.

20. (**2.5pts**) If we omit a variable that is autocorrelated, will our error term be autocorrelated? Explain your answer.

21. We have two individuals—Ali and Bob—and would like to use them to estimate the causal effect of college on income. Let $\text{Income}_{1,i}$ denote the income of individual i if she went to college and $\text{Income}_{0,i}$ if she did not.

$$\begin{aligned} \text{Income}_{1,\text{Ali}} &= 75,000 & \text{Income}_{0,\text{Ali}} &= 70,000 \\ \text{Income}_{1,\text{Bob}} &= 65,000 & \text{Income}_{0,\text{Bob}} &= 50,000 \end{aligned}$$

i. (**3pts**) Calculate the causal effect of education for Ali.

ii. (**3pts**) Calculate the causal effect of education for Bob.

iii. (**2.5pts**) Is the treatment effect constant? Briefly explain.

iv. (**3pts**) Bob went to college; Ali did not. Estimate the effect of college using the difference estimator—the difference in the mean of the treatment group (college) and the mean of the control group (no college).

v. (**4pts**) Do we have selection bias? Briefly explain your answer.

22. (**2.5pts**) What is the fundamental problem of causal inference?

23. (**4pts**) What problem does instrumental variables attempt to solve? How does it do it?

24. (**3pts**) What does it mean for an instrumental variable to be *relevant*?

25. The probability limit of the instrumental variables estimator is

$$\text{plim}\left(\hat{\beta}_1^{\text{IV}}\right) = \beta_1 + \frac{\text{Cov}(z,\, u)}{\text{Cov}(z,\, x)}$$

where z is our instrument, u is the disturbance, and x is our endogenous variable.

i. (**4pts**) How do the two requirements of a valid instrument enter into this equation?

ii. T/F (**2.5pts**) If we have a valid instrument, then $\hat{\beta}_1^{\text{IV}}$ is a consistent estimator for β_1.

26. (**2.5pts**) Consider the random walk

$$u_t = u_{t-1} + \varepsilon_t$$

where ε_t is stationary.

Take the difference between u_t and its lag. Is this difference stationary?

27. Imagine we've estimated the model

$$\text{Income}_i = \beta_0 + \beta_1 (\text{Military service})_i + u_i$$

using two-stage least squares.

Income_i gives the income of individual i, and $\text{Military service}_i$ is a binary indicator variable for whether individual i served in the military.

Our instrument for military service is Male_i, an indicator variable for whether or not the individual is male.

The following two tables give the results of the first and second stage.

First-stage results (Outcome variable: Military service)

Term	Coef. estimate	Standard error	*t* stat.	*p* Value
Intercept	0.014	0.008	1.75	0.0801
Is Male	0.120	0.046	2.61	0.0091

Second-stage results (Outcome variable: Income)

Term	Coef. estimate	Standard error	*t* stat.	*p* Value
Intercept	31,153.71	7,829.2	3.98	<0.0001
Military service	1,104.85	329.10	3.36	0.0008

Questions on the next page...

i. (**3pts**) Write out the first-stage model that we estimated.

ii. (**3pts**) Interpret the first-stage results. (What do they say about gender differences in military service?)

iii. (**3pts**) Does it appear as though we have a *relevant* instrument? Explain your answer.

iv. (**3pts**) Does it seem likely that our instrument is *exogenous*? Explain your answer.

v. (**3pts**) Assuming that we have a valid instrument, interpret the second-stage results.

Extra credit: Venn diagrams(!)

- Each circle illustrates a variable.
- Overlap refers to the (share of) correlatation between two variables.
- Dotted borders denote *omitted* variables.

Figure A

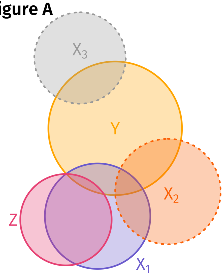


Figure B

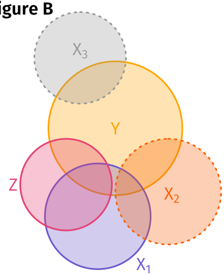


Figure C

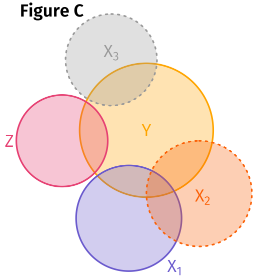
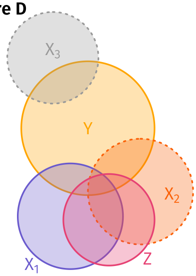


Figure D



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

EC$_1$. (**2pts**) Label the area(s) in **Figure A** that make us concerned about omitted-variable bias.

EC$_2$. (**2pts**) In which figures is Z a valid instrument for X_1?

EC$_3$. (**2pts**) In which figures is Z *relevant* for X_1?

EC$_4$. (**2pts**) In which figures is Z *exogenous* (with respect to X_1)?

EC$_5$. (**2pts**) On the back of this page, draw a Venn diagram that has two valid instruments for an endogenous variable.