Take Home Examination for ECON 6336/PPPE 7319

February 16, 2017

Your responses to this Exam are due at the beginning of class on February 23, 2017. Make sure that you provide complete responses to the questions.

1. In Freedman's 1991 article, "Statistical Models and Shoe Leather", Freedman put forth as an example of exemplary study, John Snow's investigation of the causes of Cholera. The following questions discuss Snow's investigation into the causes of Cholera.

   (a) John Snow essentially uses a comparison of water provided by the Lambeth Water Company and the Southwark and Vauxhall company as a natural experiment to investigate the causal impact of dirty water as a conduit for Cholera. First, if you could devise a randomized controlled trial to test this hypothesis? How would you do so? What issues would randomization potentially surmount? How would compliance with assignment to treatment possibly contaminate your results?

   (b) Consider Table 1 (Snow's Table IX) on page 298 of the manuscript. Explain how the data presented in the graph is consistent with the hypothesis is consistent with the hypothesis that dirty water is the conduit via which Cholera is transmitted?

   (c) What are some issues that Snow identified that possibly confound the hypothesis that dirty water is the conduit via which Cholera is transmitted?

   (d) What evidence does Snow offer to support the assumption that, when examining Lambeth and Southwark and Vauxhall water distribution, the potential confounders are not an issue?

2. In Howard Bloom's article, "Accounting for No Shows in Experimental Evaluation Designs", Bloom provides some methods to deal with the problem of No Shows. The goal is to estimate the following:

$$PE = E(Y_P) - E(Y_{P^*})$$

(a) How does the problem of No-Shows preclude one's ability to estimate the program effect despite the program effect despite their being random assignment to treatment?

(b) What parameter does random assignment to treatment allow you estimate if you elide consideration of taking up treatment; that is you ignore whether or not you took up treatment?

(c) What assumption(s) allows you to write the expectation of the control group $E(Y_C)$ as the weighted average of the mean outcomes for no-shows and the appropriate counterfactual where the weights are the no show rate?

$$E(Y_C) = (K) \times E(Y_{NS}) + (1 - K) \times E(Y_{P^*})$$

3. What is "selection?" Why is it a problem for causal analysis if individuals can choose whether or not they are treated by a given intervention?

4. Explain how randomized controlled trials overcome selection bias.

5. What is the difference between internal validity and external validity? Does the fact that a study is internally valid mean it is externally valid?

6. Bui, Craig and Imberman (2014) estimate the effect of being in a "gifted and talented" program that provides extra services for high-performing students on their academic achievement. They use a regression-discontinuity design that is based on the fact students must have a test score above a certain threshold in order to qualify for the gifted and talented program.

   (a) Explain in words how the regression discontinuity method leads to a valid estimate of the causal effect of being in a gifted and talented program on educational outcomes.

   (b) What are the assumptions under which this method is valid?

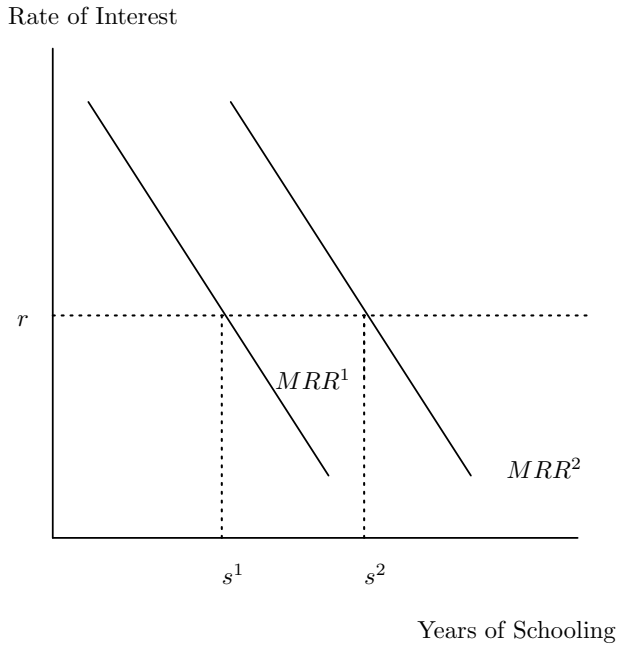
   (c) Would it be a problem for this approach if not all students who were eligible for the program decided to enroll in it?

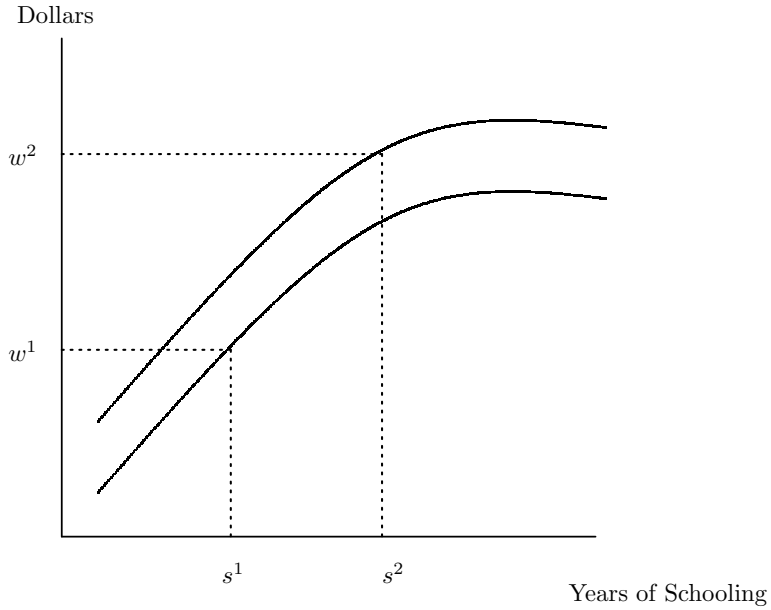
7. What are the similarities between physical and human capital? In what ways do physical and human capital differ?

8. If obtaining additional schooling always leads to some increase in earnings, why don't people enroll in school for their entire lives?

9. Walter is a chemistry teacher who earns $\$50,000$ per year, while Jesse is unemployed. Both Walter and Jesse want to go back to school to earn a business degree to help them learn how to operate their new business. The tuition each one would have to pay is the same, and they have agreed to each pay half of the tuition regardless of which one attends. Assume they cannot attend part-time. Do they indeed face the same cost of enrolling in this program? Why or why not?

10. You observe the following combination of educational choices and earnings for two individuals:


Rate of Interest
r
MRR^1
MRR^2
s^1
s^2
Years of Schooling


Dollars
w^2
w^1
s^1
s^2
Years of Schooling

Can you use those two data points to estimate the causal return to schooling? Why or why not?

11. Name a few of the limitations of economists' treatment of the returns to education that Holm identifies in his 2011 IZA article.

12. What evidence does Dr. Skimmyhorn provide in his 2016 article to support the assumption of random assignment to the boot camp? In words, explain how this evidence supports the assumption?

13. Summarize what Dr. Skimmyhorn finds in a paragraph.

14. The government of Tuvalu has hired you to estimate the returns to schooling amongst their citizens. You have administrative data on the amount of education for each person's mother (mothed) and father (fathed), parental income when in elementary school (income), IQ score (IQ) and completed education (education). You estimate a regression of wages at age 30 on these variables. Can you interpret the coefficient on completed education as the causal effect of education on wages? Why or why not?

15. Concerned over the potential problems with the regression from the above problem, you have decided to look only at identical twins. How would you use twins to estimate the returns to schooling? What problems might there be in interpreting the resulting estimate as the causal effect of schooling on wages?