Econ 300 Spring 2020 Final Exam Preparation Questions

1. To implement Sharp RD, you run a linear regression:
$Y = \alpha + \delta Above + \gamma S + \eta(S * Above) + \varepsilon$, where S is the running variable subtracted the cutoff value, Above=1 if S ≥0, Above=0 if S<0. Y is outcome. Which parameter gives the causal impact in RD?
a. δ
b. α
c. γ
d. η

2. You are studying the impact of class size on earnings. You know that the rule determining class size is 1) If there are 30 students in a grade, the class size is 30; 2) if there are 31 students in a grade, students will be split into 15, and 16. In this example, what is the running variable?
a. Class size
b. Earning
c. Number of students in a grade
d. Number of grade in a school

3. Which one of the following is **FALSE** about RD?
a. When implementing RD, linear model always gives us the best fit
b. Fuzzy RD is a type of RD
c. Sharp RD is a type of RD
d. The idea of RD is to mimic randomization around the cutoff

4. Suppose you are studying the impact of smoking on health. You know on 9/1/2009 there is an increase in Tabaco tax in California but not in other states. Which identification strategy is best fit for this analysis?
a. Differences-in-difference
b. Regression discontinuity
c. Randomized control trial
d. Simple OLS

5. Which interpretation about RD estimate is **FALSE**?
a. RD has great tests to explicitly check its identification assumption
b. RD's identification assumption will fail if there is perfect manipulation
c. RD estimate is a LATE because it estimates the average treatment effect only for people above the cutoff value
d. RD estimate is a LATE because sometimes it estimates the average treatment effect for people who are late to class.

6. Which is **NOT** a test for the identification assumption?
a. Pre-trend test
b. Covariate smoothness test
c. Balance table
d. F-test

Part II. Short questions.

1. a. Use an example to explain how you use regression discontinuity method and why you can get credible causal estimate. (You can take an example of either a Sharp or Fuzzy RD)(5pts)

b. Explain the difference between Sharp and Fuzzy RD. (5pts)

2. Suppose you are interested in the research question: does raising minimum wage lead to lower employment? You know there is a minimum wage increase for the state of New York in 2012 but not for other states.
a. Can you compare the employment rate before and after this policy change in NY to get the causal effect of the minimum wage increase on employment? Explain.(5pts)

b. Can you design a research to get the causal effect? Explain your research design. (Don't have to write the regression)(5pts)

3. Recall in the class for DID, we examined the topic whether lowering Minimum Legal Drinking Age (MLDA) leads to more death. The state of AL lowered MLDA from 21 to 19 in 1975, while the state of AR didn't. We specified a DID regression to get the causal estimate of this policy change on death rate. Below are the regression results. AL is a dummy variable whether state is AL. Post75 is a dummy whether the year is after 1975. AL_post75 is the interaction term of the two dummy variables. Death rate is defined as number of death among 100,000 people.

Source	SS	df	MS	Number of obs =	28
				F(3, 24) =	6.10
Model	5571.97376	3	1857.32459	Prob > F =	0.0031
Residual	7301.79981	24	304.241659	R-squared =	0.4328
				Adj R-squared =	0.3619
Total	12873.7736	27	476.806428	Root MSE =	17.443

mrate	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
AL	-10.80907	11.03162	-0.98	0.337	-33.57721	11.95908
post75	-35.63727	9.728975	-3.66	0.001	-55.71689	-15.55765
AL_post75	14.24673	13.75885	1.04	0.311	-14.15014	42.6436
_cons	162.5298	7.800534	20.84	0.000	146.4302	178.6293

a. What is the causal impact of lowering MLDA on death rate? Does the sign meet your expectation? (5pts)

b. Look at the standard error and t-stat of the causal estimate, what can you conclude? (5pts)

c. What does the coefficient of post75 (-35.63727) mean? (5pts)

Part III. Long questions.

1. To study whether Medicare coverage contributes to better health status, Card et al use a regression discontinuity design. People who are 65 and older are qualified for Medicare and who are under 65 are generally not qualified, with the exception that they are receiving Social Security Disability Insurance.

a. What is the running, outcome and treatment variables in this example?

b. What is the RD identification assumption in this example?

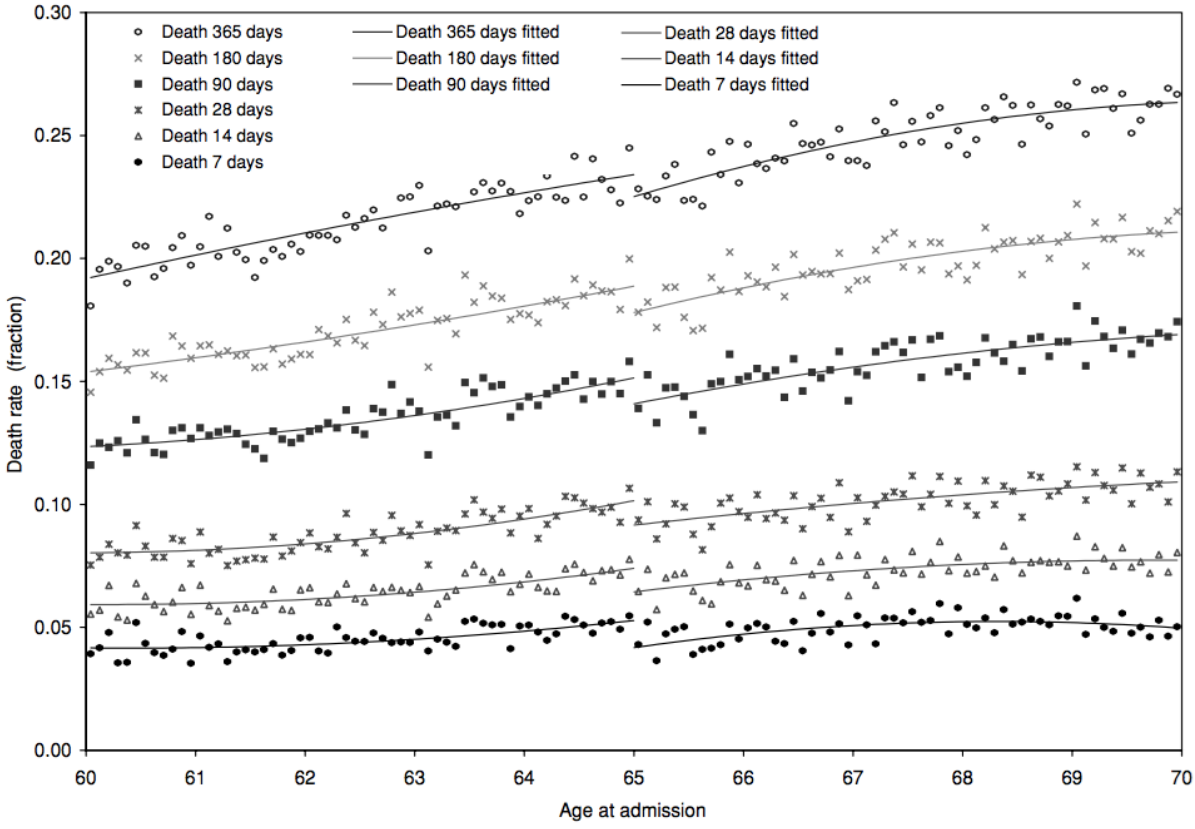
c. Draw a histogram to test the identification assumption under the condition that many 64 year olds lie about their age. (They would say they are 65 but in fact they are 64)

d. What does the histogram test in c) suggest about your identification assumption? Do you think RD is valid in this case?

d. Draw a covariate test assuming the identification assumption is valid. Pick one covariate that you see fit

f. Is this an example of Sharp or Fuzzy RD? Explain.

h. Look at the result graph below. What is your conclusion of the research?



Quiet Study Areas- Regression Discontinuity [30 pts]

In a one-time trial at a large urban university the school is offering a dedicated quiet study space for all individuals that live further than 15 miles from campus. If an individual lives 16 miles from campus they receive access to a study room in the library. If an individual lives 14 miles away from campus they do not have access to the study room. A regression discontinuity design can be used to measure the impact of having a study room on student's GPA.

1. What is the running variable?

2. Which individuals are treated?

3. What is the cutoff for treatment?

The RD regression estimates for a model with a window of within 5 units of the cutoff is

$$\widehat{gpa} = 2.7 + 0.06\, treat - 0.08\, running + 0.10 * running * treat$$

4. What is the impact of being assigned a study room on GPA? (5 points)

5. What is the assumption of a regression discontinuity design?

6. What are some ways we can check if the assumptions of RD are likely to hold?