ECO 348K, Applied Econometrics and Data Analysis **Prof. Tom Vogl**

Spring 2019 **University of Texas at Austin**

PROBLEM SET 6: DOES ACADEMIC PROBATION HELP OR HURT STUDENTS?

DUE AT 4 PM ON MONDAY 5/13

Many academic instutions place poorly-performing students on probation. A student on probation must maintain a grade-point average (GPA) above a minimum standard in the next term to avoid a one-year suspension from the university. The effect of being placed on probation on subsequent academic performance is unclear. On the one hand, probation may serve as a wake-up call and incentivize better performance. On the other hand, it may discourage already struggling students.

In this problem set, you will analyze administrative data from a large Canadian university to estimate the effect of being placed on academic probation on subsequent academic outcomes. The dataset contains all students who enrolled in the university from 1996 to 2003, with information on background characteristics, academic performance in the first two years, and graduation within 4, 5, and 6 years of entry. (Some students have missing graduation data because the study ended before their 4th, 5th, or 6th years.) At this university, if a student's GPA falls below some cutoff, the university places the student on probation. This rule gives rise to a compelling regression discontinuity design. The variable *dist_from_cut* is equal to the student's 1st-year GPA minus the probation cutoff. According to the rule, the university should place a student on academic probation after the 1st year if and only if *dist_from_cut* is negative.

Throughout your analysis, use two methods to approximate the conditional expectation function: a global 3rd-degree polynomial and a local linear regression with a bandwidth of 0.6.

1. Plot the histogram of *dist_from_cut*, allowing for a discontinuity at zero. You may choose your bin width. Do students bunch on either side of the probation cutoff? If bunching is apparent, what mechanism might explain it? If bunching is apparent, is it large or small in comparison to other observed noise in the histogram? Do you think your results represent a threat to the regression discontinuity design?
2. The last five variables in the dataset are predetermined. Test whether the conditional expectations of these variables change discontinuously at the cutoff. Do you think your results represent a threat to the regression discontinuity design?
3. Estimate the effect of falling below the cutoff on the probability of being placed on 1st-year probation. Also estimate the effect of falling below the cutoff on the probability of ever being placed on probation. Which effect is larger? Why?

4. For 1st-year probation and ever probation, create graphs that plot the local means and the polynomial fit against the running variable. You may choose your bin width for the local means. Describe how the two graphs relate to your results for question (3).
5. Estimate the effect of falling below the cutoff on medium-term outcomes: the 2nd-year GPA and the probability of dropping out of the university after the 1st year.
    - Do you think it is reasonable to interpret the discontinuity in dropout rates as the effect of actually being placed on probation after the first year? Why or why not?
    - Given that falling below the cutoff affects dropout, do you think the estimated effect on the 2nd-year GPA is unbiased?
6. Estimate the effect of falling below the cutoff on the probability of graduating within 4, 5, and 6 years.[1] Also estimate these effects separately for students who had above and below median grades in high school. Interpret your results.
7. Suppose we thought that the graduation discontinuities were driven by an effect of *ever* being placed on probation, rather than an effect of being placed on probation specifically after the first year. Use two-stage least squares to estimate how ever being placed on probation affects the probability of graduating within 4, 5, and 6 years. You do not need to perform separate analyses by high school ranking. How do your estimates relate to your results for questions (3) and (6)?

[1] There appear to be some minor coding errors in the graduation variables. Do not dwell on them.